



THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

Maximal operators and Hilbert transforms along variable non-flat homogeneous curves

Citation for published version:

Guo, S, Hickman, J, Lie, V & Roos, J 2017, 'Maximal operators and Hilbert transforms along variable non-flat homogeneous curves', *Proceedings of the London Mathematical Society*, vol. 115, no. 1, pp. 177-219. <https://doi.org/10.1112/plms.12037>

Digital Object Identifier (DOI):

[10.1112/plms.12037](https://doi.org/10.1112/plms.12037)

Link:

[Link to publication record in Edinburgh Research Explorer](#)

Document Version:

Peer reviewed version

Published In:

Proceedings of the London Mathematical Society

General rights

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.



MAXIMAL OPERATORS AND HILBERT TRANSFORMS ALONG VARIABLE NON-FLAT HOMOGENEOUS CURVES

SHAOMING GUO, JONATHAN HICKMAN, VICTOR LIE, AND JORIS ROOS

ABSTRACT. We prove that the maximal operator associated with variable homogeneous planar curves $(t, ut^\alpha)_{t \in \mathbb{R}}$, $\alpha \neq 1$ positive, is bounded on $L^p(\mathbb{R}^2)$ for each $p > 1$, under the assumption that $u : \mathbb{R}^2 \rightarrow \mathbb{R}$ is a Lipschitz function. Furthermore, we prove that the Hilbert transform associated with $(t, ut^\alpha)_{t \in \mathbb{R}}$, $\alpha \neq 1$ positive, is bounded on $L^p(\mathbb{R}^2)$ for each $p > 1$, under the assumption that $u : \mathbb{R}^2 \rightarrow \mathbb{R}$ is a measurable function and is constant in the second variable. Our proofs rely on stationary phase methods, TT^* arguments, local smoothing estimates and a pointwise estimate for taking averages along curves.

1. INTRODUCTION

This paper focuses on the study of certain maximal and singular integral operators that act by integration along *variable* homogeneous curves in the plane, of the form

$$\Gamma_u^\alpha(t) := (t, u \cdot [t]^\alpha),$$

where here and throughout the paper, the exponent α is a fixed positive real number, the notation $[t]^\alpha$ stands for either $|t|^\alpha$ or $\text{sgn}(t)|t|^\alpha$, while the “coefficient” $u(\cdot, \cdot)$ is allowed to change depending on the base point $(x, y) \in \mathbb{R}^2$.

More precisely, given $u : \mathbb{R}^2 \rightarrow \mathbb{R}$ a *measurable* function and $0 < \varepsilon_0 \leq \infty$ a parameter, we consider the following objects:

- the $(\varepsilon_0$ -truncated) **maximal operator along** Γ_u^α , defined by

$$\mathcal{M}_{u, \varepsilon_0}^{(\alpha)} f(x, y) = \sup_{0 < \varepsilon < \varepsilon_0} \frac{1}{2\varepsilon} \int_{-\varepsilon}^{\varepsilon} |f(x - t, y - u(x, y)[t]^\alpha)| dt. \quad (1.1)$$

- the $(\varepsilon_0$ -truncated) **Hilbert transform along** Γ_u^α , given by

$$\mathcal{H}_{u, \varepsilon_0}^{(\alpha)} f(x, y) = \text{p.v.} \int_{-\varepsilon_0}^{\varepsilon_0} f(x - t, y - u(x, y)[t]^\alpha) \frac{dt}{t}. \quad (1.2)$$

For convenience, in what follows we will use the convention that $[t]^1 = t$. Moreover, when $\alpha = 1$, we will leave out the dependence on α and simply write $\mathcal{M}_{u, \varepsilon_0}$ and $\mathcal{H}_{u, \varepsilon_0}$. The same principle applies to $\varepsilon_0 = \infty$: in this case we will simply write $\mathcal{M}_u^{(\alpha)}$ and $\mathcal{H}_u^{(\alpha)}$, respectively.

A difficult problem in the area of harmonic analysis is to understand the weakest possible regularity assumptions on u that guarantee the L^p boundedness of $\mathcal{M}_{u, \varepsilon_0}^{(\alpha)}$ and $\mathcal{H}_{u, \varepsilon_0}^{(\alpha)}$. Our aim in this paper is to provide a partial solution to this problem when we impose a nontrivial curvature condition by requiring $\alpha \neq 1$.

We will now state our main results. The first result regards the boundedness of the maximal operator (1.1) and extends an earlier result of Marletta and Ricci [MR98].

Date: October 18, 2016.

2010 *Mathematics Subject Classification.* 42B20, 42B25, 44A12.

The third author was supported in part by NSF grant DMS-1500958. The fourth author is supported by the German National Academic Foundation.

Theorem 1.1. *Let $\alpha > 0$ and $\alpha \neq 1$. Then the following hold:*

- (1) [MR98] *If $u : \mathbb{R}^2 \rightarrow \mathbb{R}$ is measurable, then for every $2 < p \leq \infty$ we have*

$$\|\mathcal{M}_u^{(\alpha)} f\|_p \leq C_{p,\alpha} \|f\|_p. \quad (1.3)$$

- (2) *If $u : \mathbb{R}^2 \rightarrow \mathbb{R}$ is Lipschitz, then there exists $\varepsilon_0 = \varepsilon_0(\|u\|_{\text{Lip}}) > 0$ such that for every $1 < p \leq 2$ we have*

$$\|\mathcal{M}_{u,\varepsilon_0}^{(\alpha)} f\|_p \leq C_{p,\alpha} \|f\|_p. \quad (1.4)$$

Here $C_{p,\alpha} \in (0, \infty)$ is a constant that depends only on p and α .

The second main result regards the boundedness of the Hilbert transform (1.2).

Theorem 1.2. *Let $\alpha > 0$ and $\alpha \neq 1$. Let $u : \mathbb{R}^2 \rightarrow \mathbb{R}$ be a measurable function and assume that*

$$u(x, y) = u(x, 0) \text{ for every } x, y \in \mathbb{R}. \quad (1.5)$$

Then we have that for all $1 < p < \infty$ the following holds:

$$\|\mathcal{H}_u^{(\alpha)} f\|_p \leq C_{p,\alpha} \|f\|_p. \quad (1.6)$$

The constant $C_{p,\alpha} \in (0, \infty)$ depends only on p and α .

Remark 1.1. We would first like to stress that the analogue of estimate (1.3) for the Hilbert transform $\mathcal{H}_u^{(\alpha)}$ fails for every $p \in (1, \infty)$ if we only assume u to be measurable. This follows by a straight-forward modification of Karagulyan's [Kar07] construction of a counterexample in the case $\alpha = 1$. However, if we assume u to be Lipschitz it is possible that the analogue of (1.4) for $\mathcal{H}_u^{(\alpha)}$ holds. Unfortunately, as of now, we are not able to prove or disprove this.

Remark 1.2. Notice that unlike the situation described in Theorem 1.1, in Theorem 1.2 we do not require any regularity assumptions on the function $u(x, 0)$. Next, we remark that as opposed to (1.4), the ε_0 -truncation is not present in statement (1.6). This is a direct consequence of a standard scaling argument that makes the truncation in $\mathcal{H}_{u,\varepsilon_0}^{(\alpha)}$ from (1.2) become irrelevant. Nevertheless, the one-variable assumption (1.5) should be understood as being strictly stronger than the Lipschitz assumption imposed in Theorem 1.1. Indeed, we have the following corollary of Theorem 1.1.

Corollary 1.3. *Under the same assumptions as in Theorem 1.2, we have*

$$\|\mathcal{M}_u^{(\alpha)} f\|_p \leq C_{p,\alpha} \|f\|_p \quad (1.7)$$

for all $1 < p \leq \infty$.

The proof of Corollary 1.3 is via an anisotropic scaling argument which we sketch presently. First of all, by the scaling $x \rightarrow \lambda x, y \rightarrow \lambda^\alpha y$, and a density argument, it suffices to show that

$$\|\mathcal{M}_{u,\varepsilon_0}^{(\alpha)} f\|_p \leq C_{p,\alpha} \|f\|_p, \quad (1.8)$$

for all compactly supported smooth functions f . Here ε_0 is the same as in Theorem 1.1. Now we approximate the chosen measurable function u pointwisely by a sequence of Lipschitz functions $\{u_n\}_{n \in \mathbb{N}}$ satisfying $u_n(x, y) = u_n(x, 0)$ for every $(x, y) \in \mathbb{R}^2$ and whose Lipschitz norms might grow to infinity. By changing variables $x \rightarrow x, y \rightarrow \lambda y$ in the L^p integration on the left hand side of (1.4), we obtain (1.8) with u replaced by u_n , and with a constant independent of $n \in \mathbb{N}$. In the end, we apply the dominated convergence theorem to conclude (1.8).

The curvature condition is fundamental in the proofs of both Theorem 1.1 and Theorem 1.2. Our approach relies on stationary phase methods, TT^* -arguments, local smoothing estimates and square function estimates. We speculate that the case $\alpha = 1$ will require a combination of time-frequency techniques and methods presented in the current paper, but this subject remains open for future investigation.

In the following we present the historical evolution of the subject that motivated our interest for this study. We then continue with a discussion of our main results, embedding them in the historical context and also state some further results.

1.1. Historical Background. The historical landmark that generated much of the literature discussed below is the so-called Zygmund conjecture. This long-standing open problem asks whether Lipschitz regularity of u suffices to guarantee any non-trivial L^p bounds for the maximal operator

$$\mathcal{M}_{u,\varepsilon_0}f(x,y) = \sup_{0 < \varepsilon < \varepsilon_0} \frac{1}{2\varepsilon} \int_{-\varepsilon}^{\varepsilon} |f(x-t, y-u(x,y)t)| dt, \quad (1.9)$$

provided ε_0 is small enough depending on $\|u\|_{\text{Lip}}$. A counterexample based on a construction of the Besicovitch-Kakeya set shows that we cannot expect any L^p bounds other than the trivial L^∞ bound if u is only assumed to be Hölder continuous with some exponent strictly smaller than one. The analogous question for $\mathcal{H}_{u,\varepsilon_0}$ is also widely open. For a detailed discussion of these conjectures, the interested reader is invited to consult Lacey and Li [LL10].

Notice that these two fundamental problems address the boundedness properties of (1.1) and (1.2) along Γ_u^α in the following context:

- *no regularity* above Lipschitz class is assumed;
- *no curvature* in the t parameter is present, *i.e.* $\alpha = 1$, and thus these objects have rich classes of symmetries.

Partial progress towards understanding the above open problems developed relatively slowly in the past three decades; in direct relation with the itemization above it revolved around the nature of regularity and/or suitable curvature conditions imposed on the vector field u .

Bourgain [Bou89] proved that for every real analytic function u there exists $\epsilon_0 > 0$ such that the associated maximal operator $\mathcal{M}_{u,\epsilon_0}$ is bounded on L^2 . His argument can be extended to L^p for all $p > 1$ by using a suitable interpolation argument. For smooth vector fields, Christ, Nagel, Stein and Wainger [CNSW99] proved, under some extra curvature conditions, that the associated maximal operator and singular integral operators are bounded on L^p for $p > 1$. The analogous result to that of Bourgain for singular integral operators was proved by Stein and Street [SS12]. Indeed, the result in [SS12] is far more general: in addition to the case of curves $(t, u \cdot t^\alpha)$ with u analytic and $\alpha \in \mathbb{N}$, they, in fact, consider all polynomials with analytic coefficients.

The first major breakthrough in terms of *regularity* came when Lacey and Li [LL06] brought tools from time-frequency analysis into the problem of Hilbert transforms along vector fields. To state these results, we introduce some notation.

Let $\psi_0 : \mathbb{R} \rightarrow \mathbb{R}$ be a non-negative smooth function supported on the set $[-2, -1/2] \cup [1/2, 2]$ such that for all $t \neq 0$,

$$\sum_{l \in \mathbb{Z}} \psi_l(t) = 1, \quad (1.10)$$

where $\psi_l(t) := \psi_0(2^{-l}t)$. For every $k \in \mathbb{Z}$, let $P_k^{(2)}$ denote the Littlewood-Paley projection in the y -variable corresponding to ψ_k . That is,

$$P_k^{(2)}f(x,y) := \int_{\mathbb{R}} f(x, y-\eta) \check{\psi}_k(\eta) d\eta. \quad (1.11)$$

Similarly, we define $P_k^{(1)}$. Now we are ready to state the main result of Lacey and Li.

Theorem 1.4 ([LL06]). *Let $u : \mathbb{R}^2 \rightarrow \mathbb{R}$ be an arbitrary measurable function. For every $p \geq 2$ there exists $0 < C_p < \infty$ such that the following hold:*

- *For all $k \in \mathbb{Z}$ we have*

$$\|\mathcal{H}_u P_k^{(2)}f\|_{2,\infty} \leq C_2 \|P_k^{(2)}f\|_2. \quad (1.12)$$

- *For all $p > 2$ and $k \in \mathbb{Z}$ we have*

$$\|\mathcal{H}_u P_k^{(2)}f\|_p \leq C_p \|P_k^{(2)}f\|_p. \quad (1.13)$$

A few years later, by further developing Lacey and Li's methods in [LL06] and [LL10], Bateman [Bat13] and Bateman and Thiele [BT13] proved the following result.

Theorem 1.5 ([Bat13], [BT13]). *Let $u : \mathbb{R}^2 \rightarrow \mathbb{R}$ be a measurable function satisfying*

$$u(x, y) = u(x, 0) \quad \text{a.e. } x, y \in \mathbb{R}. \quad (1.14)$$

Then for every $1 < p < \infty$ there exists $0 < C_p < \infty$ such that

$$\|\mathcal{H}_u P_k^{(2)} f\|_p \leq C_p \|P_k^{(2)} f\|_p \quad (1.15)$$

uniformly in $k \in \mathbb{Z}$. Moreover, for all $p > 3/2$, we have

$$\|\mathcal{H}_u f\|_p \leq C_p \|f\|_p. \quad (1.16)$$

An earlier result on maximal operators and Hilbert transforms along one-variable vector fields can be found in Carbery, Seeger, Wainger and Wright [CSWW99]. An especially interesting aspect of that work is that they gave an endpoint result on the product Hardy space $H_{\text{prod}}^1(\mathbb{R} \times \mathbb{R})$ under certain curvature assumption on the function u .

To pass from (1.15) to (1.16), Bateman and Thiele [BT13] relied crucially on the commutation relation

$$\mathcal{H}_u P_k^{(2)} = P_k^{(2)} \mathcal{H}_u. \quad (1.17)$$

Unfortunately, this relation fails for the maximal operator \mathcal{M}_u as it is only a sub-linear operator. For this reason the following problem still remains open.

Open Problem 1. *Let $u : \mathbb{R}^2 \rightarrow \mathbb{R}$ be a measurable function satisfying (1.14). Is \mathcal{M}_u bounded on $L^p(\mathbb{R}^2)$ for some $p < \infty$?*

This open problem served as our original motivation for the aim formulated at the beginning of our current paper. Further motivation for our study is provided by the rich literature addressing the boundedness of maximal and singular Radon transforms, focusing on curvature conditions. This corresponds to the case of Γ_u^α when $\alpha \neq 1$.

Apart from the work [MR98] that has been mentioned already, one more relevant result from this body of literature is due to Seeger and Wainger [SW03]. In this paper the variable curve $(t, u(x, y) \cdot [t]^\alpha)_{t \in \mathbb{R}}$ appears as a special case of the more general curve $\Gamma(x, y, t)$ which satisfies some convexity and doubling hypothesis uniformly in (x, y) . For such curves $\Gamma(x, y, t)$, the authors proved that the associated maximal operator and singular integral operators are bounded on L^p for $p > 1$.

For more results in the same spirit, we refer to Nagel, Stein and Wainger [NSW79], Seeger [See94] and the references therein.

1.2. Comments on the main results. We will start our discussion with Theorem 1.1.

As mentioned earlier, item (1) is the restatement of the result of Marletta and Ricci in [MR98]. To prove this result they used Bourgain's result on the circular maximal operator [Bou86] as a black-box. In contrast, we will provide an alternative approach that is more self-contained. In a certain sense we are unraveling the mechanism behind the boundedness of the circular maximal operator, by using the local smoothing estimates of Mockenhaupt, Seeger and Sogge [MSS92] and the $l^2 L^p$ decoupling inequalities for cones of Wolff [Wol00], Bourgain [Bou13] and Bourgain and Demeter [BD15]. We do not claim any originality in this approach: that local smoothing estimates can be used to prove the boundedness of the circular maximal operator has already been pointed out in [MSS92]. Moreover, the observation that decoupling inequalities for cones can provide certain progress toward the local smoothing conjecture is due to Wolff [Wol00].

The next comment regards both items (1) and (2): recall that one of the main obstacles in the analysis of the maximal operator is that the analogue of the commutation relation (1.17) fails due to sublinearity (see also (1.19) below). Thus, in order to prove estimates (1.3) and

(1.4) our strategy is to work with all frequency annuli at the same time and take advantage of the non-trivial curvature provided by $\Gamma_u^\alpha(t) = (t, u(x, 0)[t]^\alpha)$ in the situation $\alpha \neq 1$.

The method of proving Theorem 1.1 in the absence of the analogue of the commutation relation (1.17) might also provide some insight toward Open Problem 1.

We now focus our discussion on Theorem 1.2. This result can be regarded as the “curved” analogue of (1.16) from Bateman and Thiele’s Theorem 1.5. In fact, in the next subsection we will state another result that includes *the single annulus* version of both Theorem 1.2 and Theorem 1.5, corresponding to (1.15) (see Theorem 1.6 below). Moreover, in a forthcoming paper of the first, third and fourth author we will be relying in part on the ideas developed by the third author in [Lie15], [Lie16] in order to extend Theorem 1.2 to the setting of general curves (not necessarily homogeneous) obeying some suitable smoothness and curvature conditions.

Regarding the proof of Theorem 1.2, we rely on several ingredients. Following the general scheme in [BT13], we first prove a single annulus estimate

$$\|\mathcal{H}_u^{(\alpha)} P_k^{(2)} f\|_p \lesssim \|P_k^{(2)} f\|_p \quad (1.18)$$

for all $p > 1$ and then we use the commutation relation

$$\mathcal{H}_u^{(\alpha)} P_k^{(2)} = P_k^{(2)} \mathcal{H}_u^{(\alpha)} \quad (1.19)$$

to pass to a square function estimate. However, it is worth stressing here that the methods through which we achieve (1.18) and then (1.6) are quite different from the ones in [BT13]: there the authors use time-frequency techniques while in our case we rely on almost-orthogonality, stationary-phase and TT^* methods derived from the presence of curvature.

This difference is also reflected in estimate (1.6) from Theorem 1.2 where we have an improved L^p range including bounds for all p close to 1. In contrast with this, the potential range for estimate (1.16) in Theorem 1.5 to hold is $p > 4/3$. This exponent is related to the exponents in the variation norm Carleson theorem [OSTTW12]. We refer to Bateman and Thiele [BT13] for a more detailed discussion.

In order to achieve the L^p bounds for all $p > 1$ we develop a pointwise estimate for taking averages along variable curves, via the shifted (strong) Hardy-Littlewood maximal function¹. This pointwise estimate has a natural geometric interpretation: roughly speaking, it says that the averages along a thickened segment² of the curve $(t, u \cdot |t|^\alpha)$ can be pointwisely controlled, up to a small logarithmic loss, by a sum of averages taken over a number of rectangles. We remark that in the case $u \equiv 1$, our proof of Theorem 1.2 reduces to an alternative proof for the L^p boundedness of the classical singular Radon transform along the curve $(t, [t]^\alpha)$ for $p \neq 2$, which does not seem to have appeared in the literature.

1.3. Further results. Here we present two more results.

As already mentioned above, the first result encompasses the single annulus estimates corresponding to Theorem 1.2 and Theorem 1.5.

Theorem 1.6. *Let $u : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ be a measurable function with $u(x, y) = (u_1(x, y), u_2(x, y))$. We then define the Hilbert transform along the variable polynomial curve*

$$(t, u_1(x, y)t + u_2(x, y)[t]^\alpha)_{t \in \mathbb{R}}$$

by

$$\mathcal{H}_u^{(\alpha)} f(x, y) = \text{p.v.} \int_{\mathbb{R}} f(x - t, y - u_1(x, y)t - u_2(x, y)[t]^\alpha) \frac{dt}{t}. \quad (1.20)$$

Suppose now that

$$u(x, y) = u(x, 0) \quad \text{a.e. } x, y \in \mathbb{R}. \quad (1.21)$$

¹See (3.19) below.

²That is, a small neighbourhood of a segment.

Then, for each $\alpha > 0$ with $\alpha \neq 2$ and each $p > 1$, there exists $C_{\alpha,p} > 0$ such that

$$\|\mathcal{H}_u^{(\alpha)} P_k^{(2)} f\|_p \leq C_{\alpha,p} \|P_k^{(2)} f\|_p, \quad (1.22)$$

uniformly in $k \in \mathbb{Z}$, where here we recall that $P_k^{(2)}$ stands for the Littlewood-Paley projection operator in the y -variable.

Notice that, by applying a partial Fourier transform in the y -variable and using Plancherel, our theorem implies the following.

Corollary 1.7. *For each $\alpha > 0$ with $\alpha \neq 2$, we have*

$$\left\| \sup_{u_1, u_2 \in \mathbb{R}} \left| \text{p.v.} \int_{\mathbb{R}} f(x-t) e^{iu_1 t + iu_2 [t]^\alpha} \frac{dt}{t} \right| \right\|_2 \leq C_\alpha \|f\|_2, \quad (1.23)$$

with a constant C_α depending only on α .

Remark that the case $\alpha = 2$ is a deep result due to the third author [Lie09] which Corollary 1.7 does not encompass due to the quadratic modulation symmetries present if $\alpha = 2$. The third author also proved bounds for the full polynomial Carleson operator [Lie11],

$$f \mapsto \sup_P \left| \text{p.v.} \int_{\mathbb{R}} f(x-t) e^{iP(t)} \frac{dt}{t} \right|, \quad (1.24)$$

where the supremum goes over all polynomials P with real coefficients of degree less than a fixed number. The proof uses a sophisticated time-frequency approach. If α is a positive integer, (1.23) is of course a corollary of the L^2 bounds for (1.24). It is interesting however, that we can prove (1.23) essentially using only Carleson's theorem as a black box.

Theorem 1.6 is about a single annulus estimate. It can be viewed as an extension of Bateman's result [Bat13]. The proof of this result is a combination of the stationary phase method with an application of Bateman's single annulus estimate [Bat13].

Next, we will state a second single annulus estimate. It is the counterpart of Lacey and Li's result Theorem 1.4, and will provide a key insight towards proving Theorem 1.1. However, it requires a completely different proof compared with Theorem 1.6.

Theorem 1.8. *Given $\alpha > 0$ and $\alpha \neq 1$. For each measurable function $u : \mathbb{R}^2 \rightarrow \mathbb{R}$, we have*

$$\|H_u^{(\alpha)} P_k^{(2)} f\|_p \leq C_{p,\alpha} \|P_k^{(2)} f\|_p, \quad (1.25)$$

for each $p > 2$. Here $C_{p,\alpha}$ does not depend on $k \in \mathbb{Z}$.

The range $p > 2$ is sharp in the sense that Theorem 1.8 fails for $p \leq 2$. This can be seen by testing the estimate against the characteristic function of the unit ball.

The proof relies on the local smoothing estimates by Mockenhaupt, Seeger and Sogge [MSS92]. The local smoothing estimates only work for $p > 2$ which is reflected in the constraint $p > 2$ in the above theorem. It is worth mentioning that we will not need the full strength of the local smoothing estimates, but only an “ ε -amount” of them, and only for a single $p > 2$. For this reason, we are able to provide a simple and self-contained proof of the local smoothing estimates we need, via decoupling inequalities for cones in \mathbb{R}^3 . This will be the content of Appendix A. We would like to emphasize again that the approach is due to Wolff [Wol00]. Moreover, in terms of the decoupling inequalities we use for cones in \mathbb{R}^3 , we do not need the full range $2 \leq p \leq 6$ in Bourgain and Demeter [BD15], but only the range $2 \leq p \leq 4$. The decoupling inequalities for p in this range again have a simple proof. For the sake of completeness, we include it here, see Appendix B. This argument is due to Bourgain [Bou13].

Structure of the paper.

- In Section 2 we prove Theorem 1.6. This is a single annulus version of the estimate in Theorem 1.2.
- In Section 3 we prove the full Theorem 1.2. The proof will rely on a vector-valued estimate for the shifted maximal operator.
- In Section 4 we show Theorem 1.8 whose proof serves as a preparation for the corresponding proof of Theorem 1.1.
- In Section 5 we provide the proof of Theorem 1.1. Theorems 1.8 and 1.1 rely on the local smoothing estimate in Theorem A.1.
- In Appendix A we provide the proof of Theorem A.1 via decoupling inequalities for cones in \mathbb{R}^3 , following the approach of Wolff [Wol00].
- In Appendix B we provide a proof of the decoupling inequalities we need, following the approach of Bourgain [Bou13].

Notation. Throughout this paper, we will write $x \lesssim y$ to mean that there exists a constant C such that $x \leq Cy$. Here, unless otherwise specified, C may only depend on fixed parameters, which will most commonly be α and p . Further we write $x \sim y$ to mean that $x \lesssim y$ and $y \lesssim x$. Moreover, $x \approx y$ will mean $x \leq 2y$ and $y \leq 2x$. Lastly, $\mathbb{1}_E$ will always denote the characteristic function of the set E .

Acknowledgements. The first and fourth authors would like to thank Christoph Thiele for his guidance on the current project. They also thank Brian Street and Po-Lam Yung for useful discussions.

2. A SINGLE ANNULUS ESTIMATE

In this section we prove Theorem 1.6. The special case $v \equiv 0$ will later be a key ingredient in the proof of Theorem 1.2.

Dropping the dependence on u, v and α in our notation, we now set³

$$\mathcal{H}f(x, y) := \int_{\mathbb{R}} f(x - t, y - v(x)t - u(x)[t]^\alpha) \frac{dt}{t}. \quad (2.1)$$

Recall that throughout this section we always assume $\alpha \notin \{1, 2\}$. We intend to show that

$$\|\mathcal{H}P_k^{(2)}f\|_p \lesssim \|P_k^{(2)}f\|_p, \quad (2.2)$$

for each $p > 1$ and $k \in \mathbb{Z}$. The proof is a combination of the TT^* method in the spirit of Stein and Wainger [SW01], and the single annulus estimate for Hilbert transforms along one-variable vector fields by Bateman [Bat13]. Here we will need a maximally truncated version of Bateman's result.

We start the proof of (2.2). By an anisotropic scaling

$$x \rightarrow x, y \rightarrow \lambda y, \quad (2.3)$$

it suffices to prove (2.2) for $k = 0$. In the rest of this section we will always assume that $f = P_0^{(2)}f$. Furthermore, we assume without loss of generality that $u(x) > 0$ for almost every x . The case $u(x) < 0$ can be handled similarly.

Observe that

$$\mathcal{H}f(x, y) = \sum_{l \in \mathbb{Z}} \int_{\mathbb{R}} f(x - t, y - v(x)t - u(x)[t]^\alpha) \psi_l(u(x)^{1/\alpha}t) \frac{dt}{t}. \quad (2.4)$$

³Here and throughout the remainder of this text we will omit the principal value notation.

Here ψ_l is as defined in (1.10). Writing $\phi_0 = \sum_{l \leq 0} \psi_l$, we split the operator \mathcal{H} into two parts:

$$\begin{aligned} \mathcal{H}f(x, y) &= \int_{\mathbb{R}} f(x - t, y - v(x)t - u(x)[t]^\alpha) \phi_0(u(x)^{1/\alpha}t) \frac{dt}{t} \\ &\quad + \sum_{l \in \mathbb{N}} \int_{\mathbb{R}} f(x - t, y - v(x)t - u(x)[t]^\alpha) \psi_l(u(x)^{1/\alpha}t) \frac{dt}{t}. \end{aligned} \quad (2.5)$$

We bound these two terms separately in the following two subsections.

2.1. Low frequency part. Here we treat the first summand on the right hand side of (2.5). The idea is to compare it with the (maximally truncated) Hilbert transform along the one-variable vector field $(t, v(x)t)_{t \in \mathbb{R}}$ given by

$$\tilde{\mathcal{H}}^* f(x, y) := \int_{\mathbb{R}} f(x - t, y - v(x)t) \phi_0(u(x)^{1/\alpha}t) \frac{dt}{t}. \quad (2.6)$$

We want the estimate

$$\|\tilde{\mathcal{H}}^* f\|_p \lesssim \|f\|_p \quad (2.7)$$

to hold for all $p > 1$. In the case $v \equiv 0$ this follows from the boundedness of the maximally truncated Hilbert transform. For an arbitrary v it is a result essentially due to Bateman [Bat13]. For a stronger variation norm estimate, see [Guo16] by the first author. Now we look at the difference, which is given by

$$\int_{\mathbb{R}} [f(x - t, y - v(x)t - u(x)[t]^\alpha) - f(x - t, y - v(x)t)] \phi_0(u(x)^{1/\alpha}t) \frac{dt}{t}. \quad (2.8)$$

Recall that

$$f(x, y) = \int_{\mathbb{R}} f(x, y - z) \check{\psi}_0(z) dz. \quad (2.9)$$

Substituting this identity into (2.8) we obtain

$$\int_{\mathbb{R}} \int_{\mathbb{R}} f(x - t, y - v(x)t - z) [\check{\psi}_0(z - u(x)[t]^\alpha) - \check{\psi}_0(z)] \phi_0(u(x)^{1/\alpha}t) \frac{dt}{t} dz. \quad (2.10)$$

Using the key restriction $|u(x)^{1/\alpha}t| \lesssim 1$ derived from (2.10), we apply the fundamental theorem of calculus to deduce

$$|\check{\psi}_0(z - u(x)[t]^\alpha) - \check{\psi}_0(z)| \lesssim \sum_{m \in \mathbb{Z}} \frac{1}{(|m| + 1)^2} \mathbb{1}_{[m, m+1]}(z) \cdot u(x)|t|^\alpha. \quad (2.11)$$

Due to the sufficiently fast decay of $(|m| + 1)^{-2}$, we will see that the summation in m does not cause any problems. For every $m \in \mathbb{Z}$ we consider the term

$$\int_m^{m+1} \int_{\mathbb{R}} |f(x - t, y - v(x)t - z)| u(x)|t|^\alpha \phi_0(u(x)^{1/\alpha}t) \frac{dt}{|t|} dz \quad (2.12)$$

arising from applying (2.11) to (2.10) in the range $m < z < m + 1$. To bound this object in L^p we will make use of the following simple observation.

Lemma 2.1. *Let $\Gamma : \mathbb{R}^2 \rightarrow \mathbb{R}$ be a measurable function and $1 \leq p \leq \infty$. If K is a non-negative measurable function of two variables such that*

$$f \mapsto \int_{\mathbb{R}} f(x - t) K(x, t) dt \quad (2.13)$$

is bounded as an operator $L^p(\mathbb{R}) \rightarrow L^p(\mathbb{R})$ with constant C , then also

$$f \mapsto \int_{\mathbb{R}} f(x - t, y - \Gamma(x, t)) K(x, t) dt \quad (2.14)$$

is bounded as an operator $L^p(\mathbb{R}^2) \rightarrow L^p(\mathbb{R}^2)$ with constant C .

Proof. Take the L^p norm of the right hand side of (2.14). For fixed x consider the quantity

$$\left(\int_{\mathbb{R}} \left| \int_{\mathbb{R}} f(x-t, y-\Gamma(x, t)) K(x, t) dt \right|^p dy \right)^{1/p}. \quad (2.15)$$

Applying Minkowski's integral inequality we bound this by

$$\int_{\mathbb{R}} \|f(x-t, y-\Gamma(x, t))\|_{L^p(dy)} K(x, t) dt. \quad (2.16)$$

Notice that $\Gamma(x, t)$ is independent of y , hence by a simple change of variable, (2.16) is equal to

$$\int_{\mathbb{R}} \|f(x-t, y)\|_{L^p(dy)} K(x, t) dt. \quad (2.17)$$

Now take the L^p norm in x and apply the hypothesis (2.13) to $g(x) = \|f(x, \cdot)\|_p$. This concludes the proof of Lemma 2.1. \square

By Minkowski's inequality, the L^p norm of (2.12) is no greater than

$$\int_m^{m+1} \left(\int_{\mathbb{R}} \int_{\mathbb{R}} \left| \int_{\mathbb{R}} |f(x-t, y-v(x)t-z)| u(x) |t|^\alpha \phi_0(u(x)^{1/\alpha} t) \frac{dt}{|t|} dy dx \right|^p dz \right)^{1/p}. \quad (2.18)$$

Defining

$$K(x, t) := u(x) |t|^\alpha \phi_0(u(x)^{1/\alpha} t) \frac{1}{|t|} \quad (2.19)$$

we see that the operator in (2.13) is dominated by the Hardy-Littlewood maximal operator, which is bounded on $L^p(\mathbb{R})$ for all $p > 1$. Hence (2.18) is bounded by (a constant multiple of) $\|f\|_p$. This completes the proof for the first summand on the right hand side of (2.5).

2.2. High frequency part. Here we handle the second summand on the right hand side of (2.5). By the triangle inequality it suffices to prove that there exists $\gamma > 0$ such that

$$\left\| \int_{\mathbb{R}} f(x-t, y-v(x)t-u(x)[t]^\alpha) \psi_l(u(x)^{1/\alpha} t) \frac{dt}{t} \right\|_p \lesssim 2^{-\gamma l} \|f\|_p. \quad (2.20)$$

By Lemma 2.1, (2.20) holds for all $p > 1$ without the exponentially decaying factor $2^{-\gamma l}$.

Hence by interpolation it suffices to prove (2.20) for $p = 2$. This will be the goal of the present subsection.

To proceed, we apply a partial Fourier transform of the left hand side of (2.20) in the y -variable. In view of Plancherel's theorem, (2.20) for $p = 2$ is equivalent to

$$\left\| \int_{\mathbb{R}} g_\eta(x-t) e^{iv(x)\eta t + iu(x)\eta[t]^\alpha} \psi_l(u(x)^{1/\alpha} t) \frac{dt}{t} \right\|_{L^2(dx d\eta)} \lesssim 2^{-\gamma l} \|f\|_2, \quad (2.21)$$

where

$$g_\eta(x) := \int_{\mathbb{R}} e^{-i\eta y} f(x, y) dy \quad (2.22)$$

denotes the Fourier transform of the function f in its second variable. By Fubini's theorem this reduces to proving the following one-dimensional estimate:

$$\left\| \int_{\mathbb{R}} g(x-t) e^{iv(x)t + iu(x)[t]^\alpha} \psi_l(u(x)^{1/\alpha} t) \frac{dt}{t} \right\|_2 \lesssim 2^{-\gamma l} \|g\|_2 \quad (2.23)$$

for all g in $L^2(\mathbb{R})$.

Observe that the phase function contains a linear term. This amounts to providing L^2 bounds for a modulation invariant operator with a special polynomial phase.

Given the shape of the operator one might guess that one has to use the time-frequency approach implemented by the third author in [Lie09], [Lie11]. However, using crucially that $\alpha \neq 1, 2$, one can rely exclusively on TT^* arguments in order to prove (2.23). A similar idea first appeared in a slightly simpler context in [GPRY16] by Pierce, Yung, the first author and

the fourth author.

The first step to prove (2.23) is to decompose the integral inside the norm on the left hand side of (2.23) into regions where t is either positive or negative. Both parts are treated in the same way, so we only detail the estimate for the positive part. Accordingly, we denote

$$Tg(x) := \int_0^\infty g(x-t) e^{iv(x)t + iu(x)t^\alpha} \psi_l(u(x)^{1/\alpha}t) \frac{dt}{t}. \quad (2.24)$$

Then we have

$$TT^*g(y) = \int_{\mathbb{R}} (\Phi_{u(y),v(y)}^l * \tilde{\Phi}_{u(x),v(x)}^l)(y-x)g(x)dx, \quad (2.25)$$

where here we set

$$\Phi_{u,v}^l(\xi) := e^{iv\xi + iu\xi^\alpha} \frac{\psi(2^{-l}u^{1/\alpha}\xi)}{\xi} \quad \text{and} \quad \tilde{\Phi}_{u,v}^l(\xi) := \overline{\Phi_{u,v}^l(-\xi)}, \quad (2.26)$$

with $\psi(\xi) := \psi_0(\xi)\chi_{(0,\infty)}(\xi)$. The kernel of TT^* satisfies

$$\begin{aligned} & |\Phi_{u(y),v(y)}^l * \tilde{\Phi}_{u(x),v(x)}^l|(\xi) \\ &= \left| \int_{\mathbb{R}} e^{i(v(y)-v(x))\eta + iu(y)\eta^\alpha - iu(x)(\eta-\xi)^\alpha} \frac{\psi(2^{-l}u(y)^{1/\alpha}\eta)}{\eta} \frac{\psi(2^{-l}u(x)^{1/\alpha}(\eta-\xi))}{\eta-\xi} d\eta \right|. \end{aligned} \quad (2.27)$$

Let us assume for the moment that $u(x) \leq u(y)$. Denoting

$$h := \left(\frac{u(x)}{u(y)} \right)^{1/\alpha} \quad \text{and} \quad a := 2^{-l}u(x)^{1/\alpha}, \quad (2.28)$$

via the change of variables

$$2^{-l}u(y)^{1/\alpha}\eta \rightarrow \eta, \quad (2.29)$$

we see that (2.27) equals to

$$a \left| \int_{\mathbb{R}} e^{iw\eta + i2^{\alpha l}\eta^\alpha - i2^{\alpha l}(h\eta - a\xi)^\alpha} \frac{\psi(\eta)}{\eta} \frac{\psi(h\eta - a\xi)}{h\eta - a\xi} d\eta \right| \quad (2.30)$$

where w is some quantity depending on x, y and l , the value of which will be irrelevant to us. In the case $u(y) \leq u(x)$ we interchange the roles of $u(x)$ and $u(y)$.

To finish this argument we use the following oscillatory integral estimate.

Lemma 2.2. *As above, assume $\alpha \notin \{1, 2\}$ and let ψ be a smooth function supported in $[1/2, 2]$. Then, there exists $\lambda > 0$ such that for all $\xi, w \in \mathbb{R}$, $0 < h \leq 1$ and $l > 0$ we have*

$$\begin{aligned} & \left| \int_{\mathbb{R}} e^{iw\eta + i2^{\alpha l}\eta^\alpha - i2^{\alpha l}(h\eta - \xi)^\alpha} \frac{\psi(\eta)}{\eta} \frac{\psi(h\eta - \xi)}{h\eta - \xi} d\eta \right| \\ & \lesssim \mathbb{1}_{[-2^{-\lambda l}, 2^{-\lambda l}]}(\xi) + 2^{-\lambda l} \mathbb{1}_{[-2, 2]}(\xi). \end{aligned} \quad (2.31)$$

We postpone the proof to the end of this section. Using the lemma we deduce that

$$|\Phi_{u(y),v(y)}^l * \tilde{\Phi}_{u(x),v(x)}^l|(\xi) \lesssim \sum_{i=1,2} (a_i \mathbb{1}_{[-2^{-\lambda l}, 2^{-\lambda l}]}(a_i \xi) + 2^{-\lambda l} a_i \mathbb{1}_{[-2, 2]}(a_i \xi)),$$

where $a_1 := 2^{-l}u(y)^{1/\alpha}$, $a_2 := 2^{-l}u(x)^{1/\alpha}$. Therefore we have

$$\begin{aligned} |\langle TT^*g, h \rangle| &\leq \int_{\mathbb{R}} \int_{\mathbb{R}} |(\Phi_{u(y),v(y)}^l * \tilde{\Phi}_{u(x),v(x)}^l)(y-x)g(x)h(y)| dx dy \\ &\lesssim 2^{-\lambda l} \left(\int_{\mathbb{R}} \int_{\mathbb{R}} M g(x) |h(y)| dx dy + \int_{\mathbb{R}} \int_{\mathbb{R}} |g(x)| M h(y) dx dy \right) \\ &\lesssim 2^{-\lambda l} \|g\|_2 \|h\|_2. \end{aligned} \quad (2.32)$$

Here M denotes the Hardy-Littlewood maximal function and we have used its L^2 boundedness as well as the Cauchy-Schwarz inequality in the last step. This concludes the proof of (2.23).

Proof of Lemma 2.2. Denote the left hand side of (2.31) by I_ξ . First note that $I_\xi = 0$ if $|\xi| > 2$. Next, if $|\xi| \leq 2^{-\lambda l}$, then the estimate follows from the triangle inequality and so we also assume that $|\xi| > 2^{-\lambda l}$. In the following we consider only η such that the integrand in the integral defining I_ξ is not zero. This implies that $\eta, h\eta - \xi \in [1/2, 2]$. We analyze the phase function

$$Q_\xi(\eta) := w\eta + 2^{\alpha l}(\eta^\alpha - (h\eta - \xi)^\alpha). \quad (2.33)$$

Note that

$$Q_\xi''(\eta) = \alpha(\alpha - 1)2^{\alpha l}(\eta^{\alpha-2} - h^2(h\eta - \xi)^{\alpha-2}), \quad (2.34)$$

$$Q_\xi'''(\eta) = \alpha(\alpha - 1)(\alpha - 2)2^{\alpha l}(\eta^{\alpha-3} - h^3(h\eta - \xi)^{\alpha-3}), \quad (2.35)$$

and observe that the vector $X := \begin{pmatrix} Q_\xi''(\eta) \\ Q_\xi'''(\eta) \end{pmatrix}$ can be written as

$$\alpha(\alpha - 1)2^{\alpha l} \begin{pmatrix} 1 & 1 \\ (\alpha - 2)\eta^{-1} & (\alpha - 2)h(h\eta - \xi)^{-1} \end{pmatrix} \begin{pmatrix} \eta^{\alpha-2} \\ -h^2(h\eta - \xi)^{\alpha-2} \end{pmatrix}. \quad (2.36)$$

This point of the argument crucially depends on the hypothesis $\alpha \neq 2$. Denoting the 2×2 matrix in the above expression by M we calculate

$$|\det(M)| = \frac{|(\alpha - 2)\xi|}{(h\eta - \xi)\eta} \gtrsim |\xi| > 2^{-\lambda l}. \quad (2.37)$$

This allows us to estimate

$$|X| \gtrsim 2^{(\alpha-\lambda)l}. \quad (2.38)$$

Invoking van der Corput's lemma [Ste93, Chapter VIII.1] we conclude that

$$I_\xi \lesssim 2^{-(\alpha-\lambda)/3l} = 2^{-\lambda l}, \quad (2.39)$$

where in the last line above we have set $\lambda := \frac{\alpha}{4}$. \square

3. PROOF OF THEOREM 1.2

Before starting our proof, for notational simplicity, we introduce the following convention: Throughout the section we omit the dependence on u, α and simply refer to our operator as \mathcal{H} . Recalling the commutation relation⁴

$$\mathcal{H}P_k^{(2)} = P_k^{(2)}\mathcal{H}, \quad (3.1)$$

by Littlewood-Paley theory it suffices to prove that

$$\left\| \left(\sum_{k \in \mathbb{Z}} |\mathcal{H}P_k^{(2)} f|^2 \right)^{1/2} \right\|_p \lesssim \|f\|_p. \quad (3.2)$$

We stress here that the one-variable assumption (1.5) is the key fact that guarantees the commutation relation (3.1). This is the only place in this section where the one-variable assumption (1.5) is explicitly used. An implicit appearance is in the estimate (3.10) for the case $p = 2$, which is the content of the previous section.

We return to the proof of (3.2). In Section 2 we already established that

$$\|\mathcal{H}P_k^{(2)} f\|_p \lesssim \|P_k^{(2)} f\|_p$$

holds (this is the case $v \equiv 0$). Here we should note that the proof in Section 2 needs a small modification in the case when $v \equiv 0$ and $\alpha = 2$. Namely, in that situation the exponential decay estimate (2.20) is essentially a special case of a well-known result due to Stein and Wainger (see [SW01, Theorem 1]).

⁴Recall that $P_k^{(2)}$ denotes a Littlewood-Paley projection in the second variable.

Fix now $k \in \mathbb{Z}$. In view of the shape of the phase of our multiplier, we decompose our operator into a low and high frequency component respectively:

$$\begin{aligned} \mathcal{H}P_k^{(2)}f(x, y) &= \sum_{l \in \mathbb{Z}} \int_{\mathbb{R}} (P_k^{(2)}f)(x-t, y-u(x)[t]^\alpha) \psi_l(u(x)^{1/\alpha}t) \frac{dt}{t} \\ &= \left(\sum_{l \leq -k/\alpha} + \sum_{l > -k/\alpha} \right) \int_{\mathbb{R}} (P_k^{(2)}f)(x-t, y-u(x)[t]^\alpha) \psi_l(u(x)^{1/\alpha}t) \frac{dt}{t}. \end{aligned} \quad (3.3)$$

We denote

$$T_{k,0}f(x, y) := \sum_{l \leq -k/\alpha} \int_{\mathbb{R}} f(x-t, y-u(x)[t]^\alpha) \psi_l(u(x)^{1/\alpha}t) \frac{dt}{t} \quad (3.4)$$

and, for $j \geq 1$

$$T_{k,j}f(x, y) := \int_{\mathbb{R}} f(x-t, y-u(x)[t]^\alpha) \psi_{-\frac{k}{\alpha}+j}(u(x)^{1/\alpha}t) \frac{dt}{t} \quad (3.5)$$

Using the triangle inequality we obtain

$$\left\| \left(\sum_k |\mathcal{H}P_k^{(2)}f|^2 \right)^{1/2} \right\|_p \lesssim \left\| \left(\sum_k |T_{k,0}P_k^{(2)}f|^2 \right)^{1/2} \right\|_p + \sum_{j \in \mathbb{N}} \left\| \left(\sum_k |T_{k,j}P_k^{(2)}f|^2 \right)^{1/2} \right\|_p. \quad (3.6)$$

As in the previous section (see (2.8) – (2.18)) we treat $T_{k,0}P_k^{(2)}f$ as a perturbation of

$$\sum_{l \leq -k/\alpha} \int_{\mathbb{R}} P_k^{(2)}f(x-t, y) \psi_l(u(x)^{1/\alpha}t) \frac{dt}{t}. \quad (3.7)$$

This yields

$$|T_{k,0}P_k^{(2)}f| \lesssim M_S(P_k^{(2)}f) + H^*(P_k^{(2)}f). \quad (3.8)$$

Here M_S denotes the strong maximal function and H^* a maximally truncated Hilbert transform applied in the first variable. Indeed, one may deduce (3.8) using the same arguments as those of Section 2.1. The vector-valued estimates for M_S follow from the corresponding estimates for the one dimensional Hardy-Littlewood maximal function which are well-known (see Stein [Ste93, Chapter II.1]). Similarly, the vector-valued estimates for H^* follow from Cotlar's inequality and the vector-valued estimates for the Hilbert transform and the maximal function. Thus we have

$$\left\| \left(\sum_{k \in \mathbb{Z}} |T_{k,0}P_k^{(2)}f|^2 \right)^{1/2} \right\|_p \lesssim \left\| \left(\sum_{k \in \mathbb{Z}} |P_k^{(2)}f|^2 \right)^{1/2} \right\|_p \lesssim \|f\|_p \quad (3.9)$$

for all $p > 1$. This finishes the proof for the first term on the right hand side of (3.6).

To bound the second term in (3.6) we will prove that there exists $\gamma_p > 0$ such that

$$\left\| \left(\sum_{k \in \mathbb{Z}} |T_{k,j}P_k^{(2)}f|^2 \right)^{1/2} \right\|_p \lesssim 2^{-\gamma_p j} \|f\|_p. \quad (3.10)$$

For $p = 2$ this follows from (2.20) with $v \equiv 0$. Note here again that in the case $v \equiv 0$, $\alpha = 2$ the estimate (2.20) is a consequence of [SW01, Theorem 1]. Hence, by interpolation it suffices to prove

$$\left\| \left(\sum_{k \in \mathbb{Z}} |T_{k,j}P_k^{(2)}f|^2 \right)^{1/2} \right\|_p \lesssim j^4 \|f\|_p. \quad (3.11)$$

Let us first note that if p is sufficiently close to 2, then (3.11) follows immediately from an interpolation argument. To carry out this interpolation, we observe the trivial pointwise bound

$$|T_{k,j}P_k^{(2)}f| \lesssim 2^{\alpha j} M_S(P_k^{(2)}f). \quad (3.12)$$

This implies that

$$\left\| \left(\sum_{k \in \mathbb{Z}} |T_{k,j} P_k^{(2)} f|^2 \right)^{1/2} \right\|_p \lesssim 2^{\alpha j} \|f\|_p \quad (3.13)$$

for all $p > 1$. Interpolating with the bound for $p = 2$ in (3.10), we can find a positive constant $\varepsilon_0 > 0$ such that (3.10) holds true for all $p \in (2 - \varepsilon_0, 2 + \varepsilon_0)$.

Recall that our goal is to prove (3.11) for all $p > 1$. For convenience we choose to present only the case $[t]^\alpha = |t|^\alpha$. The other case $[t]^\alpha = \text{sgn}(t)|t|^\alpha$ can be treated by the same arguments.

Our strategy is to derive a sufficiently fine-grained pointwise estimate of $|T_{k,j} P_k^{(2)} f|$ by appropriate shifted maximal functions and then apply vector-valued bounds to conclude (3.11). The use of the shifted maximal operator was inspired by the work of the third author [Lie16] (see there Section 2.4., Lemma 2).

First let us consider the case $k = 0$. By definition we have

$$T_{0,j} P_0^{(2)} f(x, y) = \iint_{\mathbb{R}^2} f(x - t, y - s - u(x)|t|^\alpha) \frac{\psi_j(u(x)^{\frac{1}{\alpha}} t)}{t} \check{\psi}(s) dt ds. \quad (3.14)$$

Up to Schwartz tails in s , this is essentially an average of f over a thickened segment of a translate of the curve $(t, u(x)|t|^\alpha)$. The idea now is to cut up this thickened curve segment into pieces that are well approximated by rectangles.

Taking absolute values and using the triangle inequality we see that the previous display is

$$\lesssim \frac{1}{\lambda_{x,j}} \int_{|t| \approx \lambda_{x,j}} \int_{\mathbb{R}} |f(x - t, y - s - u(x)|t|^\alpha) \check{\psi}(s)| ds dt, \quad (3.15)$$

where $\lambda_{x,j} := 2^j u(x)^{-1/\alpha}$ and the notation $|t| \approx \lambda$ means $\frac{1}{2}\lambda \leq |t| \leq 2\lambda$. By the rapid decay of $\check{\psi}$ this is

$$\lesssim \sum_{\tau \in \mathbb{Z}} \frac{1}{(1 + |\tau|)^{10}} \frac{1}{\lambda_{x,j}} \int_{|t| \approx \lambda_{x,j}} \int_{\tau}^{\tau+1} |f(x - t, y - s - u(x)|t|^\alpha)| ds dt. \quad (3.16)$$

Once at this point, the intuition is given by the following observation: the function $f = P_0^{(2)} f$ “sees” the y -universe in unit steps; that is, f is morally y -constant on segments of length one. Consequently, it is natural to further discretize the location of t in intervals on which the variation of the term $u(x)|t|^\alpha$ does not exceed the order of one. This invites us to consider the following construction.

Set $\delta_{x,j} := 2^{-(\alpha-1)j} u(x)^{-1/\alpha}$ and cover the region $\frac{1}{2}\lambda_{x,j} \leq |t| \leq 2\lambda_{x,j}$ by intervals $\{I_m\}_{m=0}^{N_j-1}$ where

$$I_m = \left\{ t : \frac{1}{2}\lambda_{x,j} + m\delta_{x,j} \leq |t| \leq \frac{1}{2}\lambda_{x,j} + (m+1)\delta_{x,j} \right\}$$

and $N_j \in \mathbb{N}$ is such that $\frac{3}{2}\lambda_{x,j} \leq N_j \delta_{x,j} \leq 2\lambda_{x,j}$. Notice that $N_j \sim 2^{\alpha j}$ and moreover that N_j can be chosen independently of x .

With this we have

$$\begin{aligned} & \frac{1}{\lambda_{x,j}} \int_{|t| \approx \lambda_{x,j}} \int_{\tau}^{\tau+1} |f(x - t, y - s - u(x)|t|^\alpha)| ds dt \\ & \lesssim \frac{1}{N_j} \sum_{m=0}^{N_j-1} \frac{1}{|I_m|} \int_{I_m} \int_0^1 |f(x - t, y - s - \tau - u(x)|t|^\alpha)| ds dt. \end{aligned} \quad (3.17)$$

We now set

$$\mathbf{C}_m := \left\{ (t, s + \tau + u(x)|t|^\alpha) : t \in I_m, s \in [0, 1] \right\} \subset I_m \times J_m,$$

where

$$J_m = \left[\tau + u(x) \left(\frac{1}{2}\lambda_{x,j} + m\delta_{x,j} \right)^\alpha, \tau + 1 + u(x) \left(\frac{1}{2}\lambda_{x,j} + (m+1)\delta_{x,j} \right)^\alpha \right].$$

Notice that, because of our choice of $\delta_{x,j}$, the thickened curve segment \mathbf{C}_m is contained in a rectangular region of comparable area. Indeed, we have by the mean value theorem,

$$|J_m| \sim_\alpha 1 + u(x)\delta_{x,j}\lambda_{x,j}^{\alpha-1} \sim_\alpha 1.$$

Thus we further have that (3.17) is bounded by a constant multiple of

$$\frac{1}{N_j} \sum_{m=0}^{N_j-1} \frac{1}{|I_m \times J_m|} \iint_{I_m \times J_m} |f(x-t, y-s)| dt ds.$$

Given a non-negative parameter σ , we define the *shifted maximal operator* as

$$M^{(\sigma)} f(z) := \sup_{z \in I \subset \mathbb{R}} \frac{1}{|I|} \int_{I^{(\sigma)}} |f(\zeta)| d\zeta. \quad (3.18)$$

Here the supremum goes over all bounded intervals I containing z , and $I^{(\sigma)}$ denotes a shift of the interval $I = [a, b]$ given by

$$I^{(\sigma)} := [a - \sigma \cdot |I|, b - \sigma \cdot |I|] \cup [a + \sigma |I|, b + \sigma |I|].$$

Note that

$$\frac{1}{|I_m \times J_m|} \iint_{I_m \times J_m} |f(x-t, y-s)| dt ds \leq M_1^{(\sigma_m^{(1)})} M_2^{(\sigma_m^{(2)} + \tau)} f(x, y),$$

where

$$\begin{cases} \sigma_m^{(1)} &:= 2^{\alpha j-1} + m, \\ \sigma_m^{(2)} &:= c_\alpha (2^j + 2^{-(\alpha-1)j} m)^\alpha, \end{cases}$$

and c_α is a constant only depending on α and $M_1^{(n)}$ (respectively, $M_2^{(n)}$) denotes the shifted maximal operator applied in the first (respectively, second) variable. Notice that since $N_j \sim 2^{\alpha j}$ and $m < N_j$ we have that $\sigma_m^{(i)} \lesssim 2^{\alpha j}$ for $i = 1, 2$.

Altogether we have now proved that

$$|T_{0,j} P_0^{(2)} f(x, y)| \lesssim \sum_{\tau \in \mathbb{Z}} \frac{1}{(1+|\tau|)^{10}} \frac{1}{N_j} \sum_{m=0}^{N_j-1} M_1^{(\sigma_m^{(1)})} M_2^{(\sigma_m^{(2)} + \tau)} f(x, y).$$

By a scaling argument

we have that for all $k \in \mathbb{Z}$ the following holds:

$$|T_{k,j} P_k^{(2)} f(x, y)| \lesssim \sum_{\tau \in \mathbb{Z}} \frac{1}{(1+|\tau|)^{10}} \frac{1}{N_j} \sum_{m=0}^{N_j-1} M_1^{(\sigma_m^{(1)})} M_2^{(\sigma_m^{(2)} + \tau)} P_k^{(2)} f(x, y). \quad (3.19)$$

Inserting these two bounds into the left hand side of (3.11) yields

$$\left\| \left(\sum_{k \in \mathbb{Z}} \left(\sum_{\tau \in \mathbb{Z}} \frac{1}{(1+|\tau|)^{10}} \frac{1}{N_j} \sum_{m=0}^{N_j-1} M_1^{(\sigma_m^{(1)})} M_2^{(\sigma_m^{(2)} + \tau)} P_k^{(2)} f(x, y) \right)^2 \right)^{1/2} \right\|_p. \quad (3.20)$$

By the triangle inequality this is no greater than

$$\sum_{\tau \in \mathbb{Z}} \frac{1}{(1+|\tau|)^{10}} \frac{1}{N_j} \sum_{m=0}^{N_j-1} \left\| \left(\sum_{k \in \mathbb{Z}} \left(M_1^{(\sigma_m^{(1)})} M_2^{(\sigma_m^{(2)} + \tau)} P_k^{(2)} f(x, y) \right)^2 \right)^{1/2} \right\|_p. \quad (3.21)$$

Thus, to show (3.11) it suffices to prove that

$$\left\| \left(\sum_{k \in \mathbb{Z}} \left(M_1^{(\sigma_m^{(1)})} M_2^{(\sigma_m^{(2)} + \tau)} P_k^{(2)} f(x, y) \right)^2 \right)^{1/2} \right\|_p \lesssim j^4 \cdot \log^2(2+|\tau|) \|f\|_p. \quad (3.22)$$

Since $\sigma_m^{(i)} \lesssim 2^{\alpha j}$ for $i = 1, 2$, by Fubini's theorem we have that (3.22) is a consequence of the following vector-valued estimate for the one-dimensional shifted maximal operator:

$$\left\| \left(\sum_{k \in \mathbb{Z}} |M^{(n)} f_k|^2 \right)^{1/2} \right\|_p \lesssim (\log \langle n \rangle)^2 \left\| \left(\sum_{k \in \mathbb{Z}} |f_k|^2 \right)^{1/2} \right\|_p, \quad (3.23)$$

where we adopt the Japanese bracket notation $\langle n \rangle := 2 + |n|$. We give the proof of this last statement below.

Let \mathcal{D} denote the set of dyadic intervals $I = [2^k m, 2^k(m+1))$ with $k, m \in \mathbb{Z}$. In accordance with the above definition we have,

$$I^{(n)} = [2^k(m-n), 2^k(m-n+1)) \in \mathcal{D}$$

for $n \in \mathbb{Z}$. For simplicity we discuss only the dyadic variant of $M^{(n)}$, defined as

$$M^{(n)} f(x) = \sup_{x \in I \in \mathcal{D}} \frac{1}{|I|} \int_{I^{(n)}} |f(y)| dy.$$

Everything here carries over to the non-dyadic version with the constants having the same dependence on n (see Muscalu [Mus14, p. 741]).

Theorem 3.1. *For $1 < p < \infty, 1 < q \leq \infty$ we have*

$$\left\| \left(\sum_{k \in \mathbb{Z}} |M^{(n)} f_k|^q \right)^{1/q} \right\|_p \lesssim (\log \langle n \rangle)^2 \left\| \left(\sum_{k \in \mathbb{Z}} |f_k|^q \right)^{1/q} \right\|_p. \quad (3.24)$$

We did not find a reference for this result in the literature, so we choose to provide a short self-contained proof. The scalar version of this estimate involves standard Calderón-Zygmund techniques and can be found in [Mus14], where the author attributes it to Stein [Ste93].

Before starting the proof we make few more observations: firstly, notice that the endpoints $1 < p = q \leq \infty$ and $q = \infty$ of Theorem 3.1 follow immediately from the scalar version, and thus interpolation establishes the result for $1 < p \leq q < \infty$. Secondly, the exponent 2 for the log loss is only chosen for convenience; the proof actually gives a slightly better exponent.

Now the proof we present below relies on a weighted estimate in the spirit of Fefferman-Stein [FS71]:

Lemma 3.2. *Let $\omega \geq 0$ be a locally integrable function. For all $\lambda > 0$,*

$$\omega(\{x : M^{(n)} f(x) > \lambda\}) \lesssim \lambda^{-1} \left(\log \langle n \rangle \|f\|_{L^1(M^{(-n)}\omega)} + \|f\|_{L^1(M\omega)} \right), \quad (3.25)$$

where M denotes the Hardy-Littlewood maximal function.

Proof. Fix $\lambda > 0$. Let \mathcal{I} be the collection of maximal dyadic intervals I such that

$$\frac{1}{|I|} \int_I |f| > \lambda. \quad (3.26)$$

Given $I \in \mathcal{I}$ we denote by \mathcal{J}_i^I the collection of dyadic intervals J such that $|J| = 2^{-i}|I|$, $J^{(n)} \subset I$ and $\frac{1}{|J|} \int_{J^{(n)}} |f| > \lambda$. This is the i -th generation of shifted subintervals of I . We call i *large* if $2^{-i}|n| < 1$ and otherwise we call i *small*. It will be important that this depends only on n . Denote $\mathcal{J}^I = \bigcup_{i \geq 0} \mathcal{J}_i^I$. Observe that

$$\{x : M^{(n)} f(x) > \lambda\} \subset \bigcup_{I \in \mathcal{I}} \bigcup_{J \in \mathcal{J}^I} J. \quad (3.27)$$

Indeed, if x is such that $M^{(n)} f(x) > \lambda$ then there exists $J \in \mathcal{D}$ with $x \in J$ and $\frac{1}{|J|} \int_{J^{(n)}} |f| > \lambda$. By definition of \mathcal{I} there is some $I \in \mathcal{I}$ with $J^{(n)} \subset I$ and therefore $x \in J \in \mathcal{J}^I$. The crucial observation is that if $J \in \mathcal{J}_i^I$ and i is large, then J is contained in $3I$.

Thus we can estimate

$$\omega(\{x : M^{(n)}f(x) > \lambda\}) \leq \sum_{I \in \mathcal{I}} \left(\sum_{\substack{J \in \mathcal{J}_i^I \\ i \text{ small}}} \int_J \omega + \int_{3I} \omega \right). \quad (3.28)$$

For $J \in \mathcal{J}_i^I$, since $(J^{(n)})^{(-n)} = J$ we have that

$$\int_J \omega \leq \frac{1}{\lambda} \int_{J^{(n)}} |f(x)| \cdot \left(\frac{1}{|J|} \int_J \omega(y) dy \right) dx \leq \frac{1}{\lambda} \int_{J^{(n)}} |f(x)| M^{(-n)}\omega(x) dx. \quad (3.29)$$

Similarly,

$$\int_{3I} \omega \lesssim \frac{1}{\lambda} \int_I |f(x)| M\omega(x) dx. \quad (3.30)$$

Thus we can estimate (3.28) by

$$\lambda^{-1} \left(\sum_{I \in \mathcal{I}} \sum_{\substack{J \in \mathcal{J}_i^I \\ i \text{ small}}} \int_{J^{(n)}} |f| M^{(-n)}\omega + \sum_{I \in \mathcal{I}} \int_I |f| M\omega \right). \quad (3.31)$$

For fixed i the $J \in \mathcal{J}_i^I$ are disjoint and we have $\bigcup_{J \in \mathcal{J}_i^I} J^{(n)} \subset I$. Since there are about $\log \langle n \rangle$ small i , the previous display is bounded by

$$\lambda^{-1} \left(\log \langle n \rangle \|f\|_{L^1(M^{(-n)}\omega)} + \|f\|_{L^1(M\omega)} \right). \quad (3.32)$$

This finishes the proof of Lemma 3.2. \square

Now let us treat the range $p \geq q$. Lemma 3.2 says that $M^{(n)}$ is a bounded operator $L^1(\widetilde{M}^{(-n)}\omega) \rightarrow L^{1,\infty}(\omega)$ with constant $\lesssim \log \langle n \rangle$ where $\widetilde{M}^{(n)} = M^{(n)} + M$. By interpolation with the trivial L^∞ bound we get

$$\int (M^{(n)}f_k)^q \omega \lesssim \log \langle n \rangle \int |f_k|^q \widetilde{M}^{(-n)}\omega \quad (3.33)$$

for all $1 < q < \infty$. Let r be the dual exponent of p/q . We have

$$\left\| \left(\sum_{k \in \mathbb{Z}} |M^{(n)}f_k|^q \right)^{1/q} \right\|_p^q = \left\| \sum_{k \in \mathbb{Z}} |M^{(n)}f_k|^q \right\|_{p/q} = \sup_{\|\omega\|_r \leq 1} \int \sum_{k \in \mathbb{Z}} (M^{(n)}f_k)^q \omega. \quad (3.34)$$

Using the previous estimate we bound this by

$$\log \langle n \rangle \int \sum_{k \in \mathbb{Z}} |f_k|^q \widetilde{M}^{(-n)}\omega. \quad (3.35)$$

By Hölder's inequality and the scalar logarithmic bound for $M^{(n)}$ we have

$$\int \sum_{k \in \mathbb{Z}} |f_k|^q \widetilde{M}^{(-n)}\omega \leq \left\| \left(\sum_{k \in \mathbb{Z}} |f_k|^q \right)^{1/q} \right\|_p^q \|\widetilde{M}^{(-n)}\omega\|_r \lesssim (\log \langle n \rangle)^{\frac{1}{r}} \left\| \left(\sum_{k \in \mathbb{Z}} |f_k|^q \right)^{1/q} \right\|_p^q. \quad (3.36)$$

This finishes the proof of our theorem.

4. PROOF OF THEOREM 1.8

In this section we prove Theorem 1.8. The main tool we will be using is the local smoothing estimate from Theorem A.1.

We start the proof by recalling that

$$\mathcal{H}_u^{(\alpha)} f(x, y) := \int_{\mathbb{R}} f(x - t, y - u(x, y)[t]^\alpha) \frac{dt}{t}. \quad (4.1)$$

We will only present the proof of the case $\alpha > 1$; the remaining case $0 < \alpha < 1$ can be treated using the same methods and is somewhat easier. As α and u will always be fixed, we will

leave out the dependence on them in our notation and simply use T to denote $\mathcal{H}_u^{(\alpha)}$. In this section, we will prove

$$\|TP_k^{(2)}f\|_p \lesssim \|P_k^{(2)}f\|_p, \quad (4.2)$$

for all $p > 2$ and all measurable function $u : \mathbb{R}^2 \rightarrow \mathbb{R}$, with a bound independent of $k \in \mathbb{Z}$ and u . By the anisotropic scaling

$$x \rightarrow x, y \rightarrow \lambda y, \quad (4.3)$$

it suffices to prove (4.2) for $k = 0$.

In order to simplify our presentation, we introduce some notation. We let

$$z = (x, y) \quad \text{and} \quad u_z := u(x, y), \quad (4.4)$$

and set v_z to be the unique integer such that

$$2^{v_z} \leq u_z < 2^{v_z+1}. \quad (4.5)$$

For a given $k_0 \in \mathbb{Z}$, define

$$u_z^{(k_0)} := 2^{k_0-v_z} u_z. \quad (4.6)$$

Observe that $u_z^{(v_z)} = u_z$, $u_z^{(k_0)} \in [2^{k_0}, 2^{k_0+1})$ and $u_z^{(k_0)} = 2^{k_0} u_z^{(0)}$.

Denote

$$T_{k_0}f(x, y) := \int_{\mathbb{R}} f(x-t, y-u_z^{(k_0)}[t]^\alpha) \frac{dt}{t}. \quad (4.7)$$

Remark 4.1. Roughly speaking, we will bound $TP_0^{(2)}f$ by the “square function”

$$\left(\sum_{k_0 \in \mathbb{Z}} |T_{k_0}P_0^{(2)}f|^p \right)^{1/p}. \quad (4.8)$$

Here we are using an l^p sum instead of an l^2 sum. This is because p is always larger than two. At first glance, it might be a bit surprising that the term (4.8) is still a bounded operator. The proof of this fact will be achieved by applying a finer decomposition for the function f , and then seeking for enough “off-diagonal” decay via a local smoothing estimate.

We now begin our analysis by performing a dyadic decomposition of the kernel $\frac{1}{t}$ around the singularity $t = 0$. In particular, let

$$T_{k_0,l}f(x, y) := \int_{\mathbb{R}} f(x-t, y-u_z^{(k_0)}[t]^\alpha) \psi_l((u_z^{(0)})^\beta t) \frac{dt}{t}, \quad (4.9)$$

where $\beta = \frac{1}{\alpha-1}$. Recall that $\alpha > 1$ and so β is always positive. The motivation for using the factor $(u_z^{(0)})^\beta$ and the choice of β will become clear much later during the main argument (see the proof of Lemma 4.1, specifically (4.31)).

We next perform a Littlewood-Paley decomposition in the x -variable, and write

$$T_{k_0}P_0^{(2)}f = \sum_{l \in \mathbb{Z}} \sum_{k \in \mathbb{Z}} T_{k_0,l}P_k^{(1)}P_0^{(2)}f. \quad (4.10)$$

We will split the sums in $l, k \in \mathbb{Z}$ in two major cases according to the behavior of the phase of the “multiplier” (see the discussion preceding (4.11) below) associated with $T_{k_0,l}$:

- (1) **low frequency case:** $l < \max\{-k, -k_0/\alpha\}$;

In this situation the operator $T_{k_0,l}$ behaves like a one-dimensional convolution operator.

- (2) **high frequency case:** $l \geq \max\{-k, -k_0/\alpha\}$;

This case splits in two subcases:

- (a) l sits below the critical point of the phase;
- (b) l sits above the critical point of the phase.

In what follows we explain the heuristic for the above partition of our analysis.

Fix $l, k \in \mathbb{Z}$ and focus on the function $T_{k_0, l} P_k^{(1)} f$. Imagine for the moment that the function $u_z^{(k_0)}$ is a constant $u^{(k_0)}$ on the interval $[2^{k_0}, 2^{k_0+1})$. Then $T_{k_0, l} P_k^{(1)}$ becomes a convolution type operator, and hence it makes sense to speak about its multiplier as

$$\int_{\mathbb{R}} e^{i2^l t \xi + i u^{(k_0)} 2^{\alpha l} [t]^\alpha \eta} \psi_0([u_z^{(0)}]^\beta t) \frac{dt}{t}. \quad (4.11)$$

Recall that $\xi \sim 2^k$ and $\eta \sim 1$. In the situation described by item (1) either $2^l \xi \lesssim 1$ or $u^{(k_0)} 2^{\alpha l} \eta \lesssim 1$. When one of these two inequalities occurs, say $u^{(k_0)} 2^{\alpha l} \eta \lesssim 1$, then in the expression (4.9), we can view $f(x - t, y - u^{(k_0)} [t]^\alpha)$ as a perturbation of $f(x - t, y)$. This is what we meant when saying that the operator $T_{k_0, l}$ behaves like a one-dimensional operator.

Assume now we are in the situation of item (2), that is $l \geq \max\{-k, -k_0/\alpha\}$. In this case we have two possibilities:

$$k < k_0/\alpha \text{ and } k \geq k_0/\alpha. \quad (4.12)$$

In the first instance, $k < k_0/\alpha$, the phase function in (4.11) does not admit any critical point, which makes this case much easier to handle. This is the reason for which we will only focus on the latter case $k \geq k_0/\alpha$.

Now, as explained above, the cutoff between case (2a) and (2b) is indicated by the stationary phase principle. Analyzing the stationary points we deduce the requirement

$$2^{l+k} \sim 2^{\alpha l + k_0} \implies |l - \beta \cdot (k - k_0)| \lesssim 1. \quad (4.13)$$

Once at this point, we let the situation in item (2a) be defined by the conditions:

$$k \geq k_0/\alpha \text{ and } l < \beta \cdot (k - k_0). \quad (4.14)$$

while the situation in item (2b) be defined by the conditions:

$$k \geq k_0/\alpha \text{ and } l \geq \beta \cdot (k - k_0). \quad (4.15)$$

With this heuristic, based on items (1) and (2) above we split (4.10) as

$$\begin{aligned} T_{k_0} P_0^{(2)} f &= \sum_{k \in \mathbb{Z}} \sum_{l \leq \max\{-k, -k_0/\alpha\}} T_{k_0, l} P_k^{(1)} P_0^{(2)} f + \sum_{k \in \mathbb{Z}} \sum_{l > \max\{-k, -k_0/\alpha\}} T_{k_0, l} P_k^{(1)} P_0^{(2)} f \\ &=: II_{k_0} + III_{k_0}. \end{aligned} \quad (4.16)$$

These two terms are treated separately in the following two subsections.

4.1. The high frequency case. In this subsection, we will treat the term III_{k_0} which is, in fact, the main term in (4.16).

Recall from our heuristic that in the situation $k < k_0/\alpha$ the phase function in (4.11) does not admit a critical point. That makes this case easier to handle. The precise arguments are the same as for the main term, so we will not detail them here. That is, we will only treat the case $k \geq k_0/\alpha$. Accordingly we redefine

$$III_{k_0} := \sum_{k \geq k_0/\alpha} \sum_{l > -k_0/\alpha} T_{k_0, l} P_k^{(1)} P_0^{(2)} f. \quad (4.17)$$

As $\alpha > 1$, in this situation we always have

$$-k_0/\alpha \leq \beta \cdot (k - k_0).$$

Thus we split III_{k_0} into two parts

$$\begin{aligned} III_{k_0} &= \sum_{k \geq \frac{k_0}{\alpha}} \sum_{l > -k_0/\alpha}^{\beta \cdot (k - k_0)} T_{k_0, l} P_k^{(1)} P_0^{(2)} f + \sum_{k \geq \frac{k_0}{\alpha}} \sum_{l > \beta \cdot (k - k_0)} T_{k_0, l} P_k^{(1)} P_0^{(2)} f \\ &=: III_{k_0}^{(1)} + III_{k_0}^{(2)}. \end{aligned} \quad (4.18)$$

Our goal here will be to prove that

$$\left\| \left(\sum_{k_0 \in \mathbb{Z}} |II_{k_0}^{(j)}|^p \right)^{1/p} \right\|_p \lesssim \|f\|_p \quad \text{for } j \in \{1, 2\}. \quad (4.19)$$

For this, we first apply the change of variables⁵ $l \rightarrow l + \beta(k - k_0)$, and use Fubini to deduce that

$$II_{k_0}^{(1)} = \sum_{l=-\infty}^0 \sum_{k > \frac{k_0}{\alpha} - \frac{l}{\beta}} T_{k_0, \beta(k-k_0)+l} P_k^{(1)} P_0^{(2)} f, \quad (4.20)$$

and

$$II_{k_0}^{(2)} = \sum_{l=1}^{\infty} \sum_{k \geq \frac{k_0}{\alpha}} T_{k_0, \beta(k-k_0)+l} P_k^{(1)} P_0^{(2)} f. \quad (4.21)$$

In order to prove (4.19) it will be sufficient to show that for all $p > 2$ we have

$$\left\| \left(\sum_{k_0 \in \mathbb{Z}} \left| \sum_{k > \frac{k_0}{\alpha} - \frac{l}{\beta}} T_{k_0, \beta(k-k_0)+l} P_k^{(1)} P_0^{(2)} f \right|^p \right)^{1/p} \right\|_p \lesssim 2^{-\gamma_p |l|} \|f\|_p, \quad (4.22)$$

and

$$\left\| \left(\sum_{k_0 \in \mathbb{Z}} \left| \sum_{k \geq \frac{k_0}{\alpha}} T_{k_0, \beta(k-k_0)+l} P_k^{(1)} P_0^{(2)} f \right|^p \right)^{1/p} \right\|_p \lesssim 2^{-\gamma_p l} \|f\|_p, \quad (4.23)$$

for some $\gamma_p > 0$. Here and throughout the paper γ_p is a positive constant that is allowed to change from line to line. We claim that for $p > 2$ there exists $\gamma_p > 0$ such that

- if $l \leq 0$ and $k > \frac{k_0}{\alpha} - \frac{l}{\beta}$, then

$$\|T_{k_0, \beta(k-k_0)+l} P_k^{(1)} P_0^{(2)} f\|_p \lesssim 2^{-\gamma_p \cdot (\alpha\beta k - \beta k_0 + l)} \|P_k^{(1)} P_0^{(2)} f\|_p, \quad \text{and} \quad (4.24)$$

- if $l > 0$ and $k \geq \frac{k_0}{\alpha}$, then

$$\|T_{k_0, \beta(k-k_0)+l} P_k^{(1)} P_0^{(2)} f\|_p \lesssim 2^{-\gamma_p \cdot (\alpha\beta k - \beta k_0 + \alpha l)} \|P_k^{(1)} P_0^{(2)} f\|_p. \quad (4.25)$$

Before we prove this claim, we will demonstrate how it is used to show (4.22) and (4.23). Here we only prove (4.22); the estimate (4.23) follows in essentially the same way.

For a fixed $l \leq 0$, we expand the L^p norm on the left hand side of (4.22), and then apply (4.24) to obtain

$$\left(\sum_{k_0 \in \mathbb{Z}} \left(\sum_{k > \frac{k_0}{\alpha} - \frac{l}{\beta}} 2^{-\gamma_p \cdot (\alpha\beta k - \beta k_0 + l)} \|P_k^{(1)} P_0^{(2)} f\|_p \right)^p \right)^{\frac{1}{p}}.$$

By applying Hölder's inequality to the summation in k , we obtain that the above expression is bounded by

$$\left(\sum_{k_0 \in \mathbb{Z}} \sum_{k > \frac{k_0}{\alpha} - \frac{l}{\beta}} 2^{-\gamma_p \cdot (\alpha\beta k - \beta k_0 + l)} \|P_k^{(1)} P_0^{(2)} f\|_p^p \right)^{\frac{1}{p}}.$$

Now we apply Fubini's theorem and exchange the order of summations in k and k_0 to further bound this by

$$2^{\gamma_p \cdot l(\alpha-1)} \left(\sum_{k \in \mathbb{Z}} \|P_k^{(1)} P_0^{(2)} f\|_p^p \right)^{\frac{1}{p}} \lesssim 2^{\gamma_p \cdot l(\alpha-1)} \|f\|_p.$$

This finishes the proof of the desired estimate (4.22).

⁵Note that $\beta(k - k_0)$ might not be an integer but this is irrelevant.

It remains to prove (4.24) and (4.25). The idea is to reduce to the local smoothing estimate provided in Theorem A.1. We will do this in a slightly more general setting in view of further applications in Section 5. To this end, given parameters $u > 0, w, l \in \mathbb{Z}$, we introduce

$$A_{u,w,l}f(z) := \int f(x-t, y-u[t]^\alpha) \psi_l(2^{\frac{w}{\alpha}}(2^{-v}u)^\beta t) \frac{dt}{|t|},$$

where $v \in \mathbb{Z}$ is such that $u \in [2^v, 2^{v+1})$. Note that $2^{-v}u \in [1, 2)$. With this notation we have $T_{k_0,l} = A_{2^{k_0}u_z^{(0)},0,l}$. The extra parameter w will only be needed in the Section 5.

Lemma 4.1. *Let $r = r_z : \mathbb{R}^2 \rightarrow [1, 2)$ be measurable and $m, k, v, l \in \mathbb{Z}$ such that*

$$M := \max\{k + l - \frac{w}{\alpha}, m + l\alpha - w + v\} \geq 0. \quad (4.26)$$

Then, for every $p > 2$ there exists $\gamma_p > 0$ such that

$$\|A_{2^v r_z, w, l} P_k^{(1)} P_m^{(2)} f(z)\|_{L_z^p} \lesssim_p 2^{-\gamma_p M} \|f\|_p, \quad (4.27)$$

where the implicit constant depends only on p and α .

Before we proceed to proving this statement let us first note that it directly implies (4.24) and (4.25) (here we use the identity $\beta + 1 = \alpha\beta$ and $\alpha > 1$).

Proof of Lemma 4.1. We have

$$A_{2^v r_z, w, l} P_k^{(1)} P_m^{(2)} f(z) = \int_{\mathbb{R}} P_k^{(1)} P_m^{(2)} f(x-t, y-2^v r_z[t]^\alpha) \psi_l(2^{\frac{w}{\alpha}} r_z^\beta t) \frac{dt}{|t|}. \quad (4.28)$$

By a change of variables $t \rightarrow 2^{l-\frac{w}{\alpha}} t$ we obtain

$$\int_{\mathbb{R}} P_k^{(1)} P_m^{(2)} f(x-2^{l-\frac{w}{\alpha}} t, y-2^{l\alpha-w+v} r_z[t]^\alpha) \psi_0(r_z^\beta t) \frac{dt}{|t|}. \quad (4.29)$$

Define $D_{a,b}f(x, y) := f(2^a x, 2^b y)$ and

$$Bf(z) := \int_{\mathbb{R}} f(x-t, y-r_z[t]^\alpha) \psi_0(r_z^\beta t) \frac{dt}{|t|}. \quad (4.30)$$

Changing variables $t \rightarrow r_z^{-\beta} t$ and using the identity $\beta = \alpha\beta - 1$ we see that B can be written in terms of the averaging operator from (A.1):

$$Bf(z) = \int_{\mathbb{R}} f(x-r_z^{-\beta} t, y-r_z^{-\beta}[t]^\alpha) \psi_0(t) \frac{dt}{|t|} = A_{r_z^{-\beta}} f(z). \quad (4.31)$$

From (4.28) – (4.30) we see that

$$A_{2^v r_z, w, l} P_k^{(1)} P_m^{(2)} f(z) = D_{-l+\frac{w}{\alpha}, -l\alpha+w-v} B D_{l-\frac{w}{\alpha}, l\alpha-w+v} P_k^{(1)} P_m^{(2)} f(z). \quad (4.32)$$

Since $D_{a,b} P_k^{(1)} P_m^{(2)} = P_{k+a}^{(1)} P_{m+b}^{(2)} D_{a,b}$, the right hand side in the previous display can be written as

$$D_{-l+\frac{w}{\alpha}, -l\alpha+w-v} B P_{k+l-\frac{w}{\alpha}}^{(1)} P_{m+l\alpha-w+v}^{(2)} D_{l-\frac{w}{\alpha}, l\alpha-w+v} f(z). \quad (4.33)$$

Since conjugation by $D_{a,b}$ is an L^p isometry, the L^p norm of (4.33) equals

$$\|B P_{k+l-\frac{w}{\alpha}}^{(1)} P_{m+l\alpha-w+v}^{(2)} f\|_p. \quad (4.34)$$

The claim now follows from Theorem A.1 once we notice that the frequency support of $P_{k+l-\frac{w}{\alpha}}^{(1)} P_{m+l\alpha-w+v}^{(2)} f$ is contained in the annulus $\|(\xi, \eta)\| \sim 2^M$. \square

4.2. The low frequency case. In this subsection, we bound the first term in (4.16). We write it as

$$\begin{aligned}
I_{k_0} &= \sum_{k \in \mathbb{Z}} \sum_{l \leq \max\{-k, -k_0/\alpha\}} T_{k_0, l} P_k^{(1)} P_0^{(2)} f \\
&= \sum_{k \in \mathbb{Z}} \sum_{l \leq -\frac{k_0}{\alpha}} T_{k_0, l} P_k^{(1)} P_0^{(2)} f + \sum_{k \leq \frac{k_0}{\alpha}} \sum_{-\frac{k_0}{\alpha} \leq l \leq -k} T_{k_0, l} P_k^{(1)} P_0^{(2)} f \\
&=: I_{k_0}^{(1)} + I_{k_0}^{(2)}.
\end{aligned} \tag{4.35}$$

The first term can be bounded by the strong maximal function and the maximally truncated Hilbert transform in the x -variable. Indeed, comparing $I_{k_0}^{(1)}$ with

$$\sum_{l \leq -\frac{k_0}{\alpha}} \int_{\mathbb{R}} P_0^{(2)} f(x - t, y) \psi_l(t) \frac{dt}{t},$$

we find that their difference is bounded by the strong maximal function. This follows by the same argument as in Section 2.1.

We pass now to the treatment of the second term, $I_{k_0}^{(2)}$. For this purpose we first define the “one-dimensional” operator

$$U_{k_0, l} f(x, y) := \int_{\mathbb{R}} f(x, y - u_z^{(k_0)}[t]^\alpha) \psi_l([u_z^{(0)}]^\beta t) \frac{dt}{t}. \tag{4.36}$$

Notice that when $[t]^\alpha$ is an even function this operator is identically zero. We next rewrite the second term as

$$I_{k_0}^{(2)} = \sum_{l=0}^{\infty} \sum_{k \leq \frac{k_0}{\alpha} - l} (T_{k_0, -\frac{k_0}{\alpha} + l} P_k^{(1)} P_0^{(2)} f - U_{k_0, -\frac{k_0}{\alpha} + l} P_k^{(1)} P_0^{(2)} f) + \sum_{l=0}^{\infty} \sum_{k \leq \frac{k_0}{\alpha} - l} U_{k_0, -\frac{k_0}{\alpha} + l} P_k^{(1)} P_0^{(2)} f. \tag{4.37}$$

For the contribution coming from the latter term, by the triangle inequality, it suffices to prove that

$$\left\| \sup_{k_0 \in \mathbb{Z}} \left| \sum_{k \leq \frac{k_0}{\alpha} - l_0} \int_{\mathbb{R}} (P_k^{(1)} P_0^{(2)} f)(x, y - u_z^{(k_0)}[t]^\alpha) \psi_{-\frac{k_0}{\alpha} + l_0}([u_z^{(0)}]^\beta t) \frac{dt}{t} \right| \right\|_p \lesssim 2^{-\gamma_p l_0} \|f\|_p, \tag{4.38}$$

for each $l_0 \in \mathbb{N}$, and for a constant γ_p depending only on p . This further follows from the pointwise bound

$$\left| \sum_{k \leq \frac{k_0}{\alpha} - l_0} \int_{\mathbb{R}} (P_k^{(1)} P_0^{(2)} f)(x, y - u_z^{(k_0)}[t]^\alpha) \psi_{-\frac{k_0}{\alpha} + l_0}([u_z^{(0)}]^\beta t) \frac{dt}{t} \right| \lesssim 2^{-\gamma_p l_0} M_S f(x, y), \tag{4.39}$$

which is again a consequence of the mean zero property of $\psi_{-\frac{k_0}{\alpha} + l_0}(t) \cdot \frac{1}{t}$. Indeed, (4.39) follows from classical Calderón-Zygmund theory. We leave the details for the interested reader.

Turning our attention towards the contribution from the former term on the right hand side of (4.37), we notice that it is enough to show that

$$\left\| \sup_{k_0 \in \mathbb{Z}} \left| \sum_{k \leq \frac{k_0}{\alpha} - l_0} (T_{k_0, -\frac{k_0}{\alpha} + l_0} P_k^{(1)} P_0^{(2)} f - U_{k_0, -\frac{k_0}{\alpha} + l_0} P_k^{(1)} P_0^{(2)} f) \right| \right\|_p \lesssim 2^{-\gamma_p l_0} \|f\|_p, \tag{4.40}$$

for each $l_0 \in \mathbb{N}$. This in turn follows from

Claim 4.2. *In the above setting, for each $l_0, j_0 \in \mathbb{N}$ and $k_0 \in \mathbb{Z}$ the following estimate holds uniformly*

$$\left\| \sup_{k_0 \in \mathbb{Z}} \left| T_{k_0, -\frac{k_0}{\alpha} + l_0} P_{\frac{k_0}{\alpha} - l_0 - j_0}^{(1)} P_0^{(2)} f - U_{k_0, -\frac{k_0}{\alpha} + l} P_{\frac{k_0}{\alpha} - l_0 - j_0}^{(1)} P_0^{(2)} f \right| \right\|_p \lesssim 2^{-\gamma_p \cdot \max\{l_0, \frac{j_0}{9}\}} \|f\|_p. \quad (4.41)$$

Proof. We first look at the case $j_0 \leq 9 \cdot l_0$. Based on the treatment of $I_{k_0}^{(2)}$, it is enough to show that

$$\|T_{k_0, -\frac{k_0}{\alpha} + l_0} P_{\frac{k_0}{\alpha} - l_0 - j_0}^{(1)} P_0^{(2)} f\|_p \lesssim 2^{-\gamma_p \cdot l_0} \|f\|_p, \quad (4.42)$$

for each fixed k_0, l_0 and j_0 . This follows from Lemma 4.1.

Next we consider the case $j_0 \geq 9 \cdot l_0$. It suffices to prove the pointwise bound

$$|T_{k_0, -\frac{k_0}{\alpha} + l_0} P_{\frac{k_0}{\alpha} - l_0 - j_0}^{(1)} P_0^{(2)} f - U_{k_0, -\frac{k_0}{\alpha} + l} P_{\frac{k_0}{\alpha} - l_0 - j_0}^{(1)} P_0^{(2)} f| \lesssim 2^{-\gamma \cdot j_0} \cdot M_S(f), \quad (4.43)$$

for some positive $\gamma > 0$. By scaling it suffices to look at the case $k_0 = 0$. We now observe that the left hand side of (4.43) can be written as

$$\left| \int_{\mathbb{R}} \left[(P_{-l_0-j_0}^{(1)} P_0^{(2)} f)(x-t, y-u_z^{(0)}[t]^\alpha) - (P_{-l_0-j_0}^{(1)} P_0^{(2)} f)(x, y-u_z^{(0)}[t]^\alpha) \right] \psi_{l_0}([u_z^{(0)}]^\beta t) \frac{dt}{t} \right|. \quad (4.44)$$

Now our claim follows by applying the fundamental theorem of calculus in the first variable of $P_{-l_0-j_0}^{(1)} P_0^{(2)} f$. \square

5. PROOF OF THEOREM 1.1

The proof is organized as follows:

- In Section 5.1 we reduce our proof to the exponential decay estimate (5.6).
- In Section 5.2 we prove this decay estimate for $p > 2$, by only assuming u to be a measurable function. This recovers the result of Marletta and Ricci [MR98].
- In Section 5.3 we prove the estimate (5.6) for $p \leq 2$. This relies on the Lipschitz assumption of the function u , a condition that is used when applying a suitable change of variables (see (5.31)). To enable that change of variables, we will introduce several auxiliary functions and make use of Lemma 5.1 on the boundedness of maximal operators along curves in lacunary directions. The proof of this lemma is provided in Section 5.4.

5.1. Preliminaries. Since we are dealing with a positive operator we may assume without loss of generality that $f \geq 0$. Furthermore, we may assume that $u(x, y) > 0$ for all $(x, y) \in \mathbb{R}^2$. We will adopt the notation (4.4) – (4.6) from the previous section.

For a positive real number u and $l \in \mathbb{Z}$ we define

$$A_{u,l} f(z) = \int_{\mathbb{R}} f(x-t, y-u[t]^\alpha) \psi_l(2^{\frac{v}{\alpha}} (2^{-v} u)^\beta t) \frac{dt}{|t|},$$

where $v \in \mathbb{Z}$ is such that $u \in [2^v, 2^{v+1})$ and $\beta := \frac{1}{\alpha-1}$. In particular, $2^{-v} u \in [1, 2)$. Observe that

$$\mathcal{M}_{u, \varepsilon_0}^{(\alpha)} f(z) \lesssim \sup_{l: v_z \in E_l} A_{u_z, l} f(z),$$

where

$$E_l := \{v \in \mathbb{Z} : 2^l \leq c_\alpha \varepsilon_0 2^{\frac{v}{\alpha}}\}, \quad (5.1)$$

for $v \in \mathbb{Z}$ and c_α is a fixed constant depending only on α . Linearizing the supremum we introduce the operator

$$\mathcal{M}f(z) := A_{u_z, l_z} f(z),$$

where $z \mapsto l_z$ is an arbitrary measurable map $\mathbb{R}^2 \rightarrow \mathbb{Z}$ such that $v_z \in E_{l_z}$ for all z . To prove our theorem it suffices to show that

$$\|\mathcal{M}f\|_p \lesssim_{\alpha, p} \|f\|_p$$

for all $1 < p < \infty$. By the Fourier inversion formula we have

$$A_{u_z, l} f(z) = \int_{\mathbb{R}^2} \widehat{f}(\xi, \eta) e^{ix\xi + iy\eta} m_l(z, \xi, \eta) d(\xi, \eta),$$

with the symbol $m_l(z, \xi, \eta)$ given by

$$m_l(z, \xi, \eta) := \int_{\mathbb{R}} e^{-i(\xi t + \eta u_z[t]^\alpha)} \psi_l(2^{\frac{v_z}{\alpha}} (u_z^{(0)})^\beta t) \frac{dt}{|t|}, \quad (5.2)$$

where $u_z^{(0)} := 2^{-v_z} u_z$, as defined in (4.5). As in our previous analysis, our approach relies on decomposing our operator relative to the behavior of the phase function of the multiplier. Our initial focus will be on the behavior of the η component of the phase which corresponds in the spatial variable to the component containing the vector field u . Thus, we will discuss the following two cases:

- the η -low frequency regime: $|\eta|2^{\alpha l} \leq 1$;
- the η -high frequency regime: $|\eta|2^{\alpha l} > 1$.

Following the above description, we perform a Littlewood-Paley decomposition in the y -variable and write

$$A_{u_z, l_z} = \sum_{k \in \mathbb{Z}} A_{u_z, l_z} P_k^{(2)} = \sum_{k \leq -\alpha l_z} A_{u_z, l_z} P_k^{(2)} + \sum_{k > -\alpha l_z} A_{u_z, l_z} P_k^{(2)}. \quad (5.3)$$

Case 1. η -low frequency.

The sum over $k \leq -\alpha l_z$ can be estimated by the strong maximal operator and is therefore bounded on all L^p , $1 < p \leq \infty$. This follows by the same arguments as in Section 2.1.

Case 2. η -high frequency.

The remaining term equals⁶

$$\sum_{k \in \mathbb{N}} A_{u_z, l_z} P_{-\alpha l_z + k}^{(2)} = \sum_{k \in \mathbb{N}} \sum_{m \in \mathbb{Z}} A_{u_z, l_z} P_m^{(1)} P_{-\alpha l_z + k}^{(2)}. \quad (5.4)$$

Fix a $k \in \mathbb{N}$, and for the following heuristic, let us imagine for a moment that u_z is a constant. Then from (5.2), the symbol of $A_{u_z, l_z} P_{-\alpha l_z + k}^{(2)}$ becomes

$$\int_{\mathbb{R}} e^{i(2^{l_z} u_z^{-\frac{1}{\alpha}} \xi t + \eta 2^{\alpha l_z} [t]^\alpha)} \psi_0(t) \frac{dt}{|t|}.$$

From the stationary phase principle it is plausible to expect an exponential decay for the L^p norm of the term (5.4) in terms of $k \in \mathbb{N}$. Indeed, this is the case for Hilbert transforms along one-variable curves (see the estimate (3.10)). However, in the present situation we do not know how to exhibit such an exponential decay. To remedy this we first remove from the term (5.4) the ξ -low frequency component; the remaining part then admits an exponential decay estimate in $k \in \mathbb{N}$.

We split the term into (5.4) in two components corresponding to

- the ξ -low frequency regime: $|2^{l_z} u_z^{-\frac{1}{\alpha}} \xi| < 1$;
- the ξ -high frequency regime: $|2^{l_z} u_z^{-\frac{1}{\alpha}} \xi| \geq 1$.

That is, we write (5.4) as

$$\sum_{k \in \mathbb{N}} A_{u_z, l_z} \sum_{m \leq -l_z + \frac{v_z}{\alpha}} P_m^{(1)} P_{-\alpha l_z + k}^{(2)} + \sum_{k \in \mathbb{N}} \sum_{m \in \mathbb{N}} A_{u_z, l_z} P_{-l_z + \frac{v_z}{\alpha} + m}^{(1)} P_{-\alpha l_z + k}^{(2)}. \quad (5.5)$$

⁶Here and in the following our notation will ignore the issue that $-\alpha l_z$ might not be an integer as this can be easily addressed by setting $P_\alpha^{(i)} = P_{\lfloor \alpha \rfloor}^{(i)}$.

Case 2.1. ξ -low frequency.

This corresponds to the first term in (5.5). The contribution from this term can be controlled by the strong maximal function using the same argument as in the η -low frequency case. We omit the details.

Case 2.2. ξ -high frequency.

To handle the latter term on the right hand side of (5.5), it suffices to show that for every $p > 1$ there exists $\gamma_p > 0$ such that

$$\left\| \sum_{m \in \mathbb{N}} A_{u_z, l_z} P_{-l_z + \frac{v_z}{\alpha} + m}^{(1)} P_{-\alpha l_z + k}^{(2)} f \right\|_p \lesssim 2^{-\gamma_p k} \|f\|_p. \quad (5.6)$$

In view of the Littlewood-Paley square function estimate, the argument splits naturally into two cases: $p > 2$ and $p \leq 2$.

5.2. The case $p > 2$. In this section we reprove the result of Marletta and Ricci [MR98] and, in particular, we only require the function u to be measurable. We also do not need to consider a truncated maximal function; under the measurability assumption, the truncated case is equivalent to the untruncated case, by a scaling argument. Hence we take $\varepsilon_0 = \infty$.

We first remark that the left hand side of (5.6) is bounded from above by

$$\left\| \left(\sum_{l \in \mathbb{Z}} \left| \sum_{m \in \mathbb{N}} A_{u_z, l} P_{-l + \frac{v_z}{\alpha} + m}^{(1)} P_{-\alpha l + k}^{(2)} f \right|^p \right)^{1/p} \right\|_p. \quad (5.7)$$

Hence, it suffices to prove that for every $l \in \mathbb{Z}$ one has

$$\left\| \sum_{m \in \mathbb{N}} A_{u_z, l} P_{-l + \frac{v_z}{\alpha} + m}^{(1)} P_{-\alpha l + k}^{(2)} f \right\|_p \lesssim 2^{-\gamma_p k} \|f\|_p. \quad (5.8)$$

For a fixed $v \in \mathbb{Z}$, recall the definition of $u_z^{(v)}$ from (4.6). Then the left hand side of (5.8) can be bounded by

$$\left\| \left(\sum_{v \in \mathbb{Z}} \left| \sum_{m \in \mathbb{N}} A_{u_z^{(v)}, l} P_{-l + \frac{v_z}{\alpha} + m}^{(1)} P_{-\alpha l + k}^{(2)} f \right|^p \right)^{1/p} \right\|_p. \quad (5.9)$$

Taking in account (5.7)–(5.9), we see that it suffices to show that for every $p > 2$ there exists $\gamma_p > 0$ such that for every $k, m \in \mathbb{N}$ and $l, v \in \mathbb{Z}$ we have

$$\left\| A_{u_z^{(v)}, l} P_{-l + \frac{v_z}{\alpha} + m}^{(1)} P_{-\alpha l + k}^{(2)} f(z) \right\|_p \lesssim 2^{-\gamma_p \max\{k, m\}} \|f\|_p. \quad (5.10)$$

Noting that $A_{u_z^{(v)}, l} = A_{2^v u_z^{(0)}, v, l}$ we recognize this as precisely the conclusion of Lemma 4.1.

5.3. L^p estimates for $p \leq 2$. In this subsection we will prove (5.6) for all $1 < p \leq 2$. Note that we are working here with two extra assumptions, as compared with the previous subsection:

- First, we assume that u is Lipschitz. This is because if u is only assumed to be measurable, then it is possible for $\mathcal{M}_{u, \varepsilon_0}^{(\alpha)}$ to be unbounded for each $p \leq 2$. One may verify this by simply plugging the characteristic function of the unit ball into the inequality.
- Second, the truncation $\varepsilon \leq \varepsilon_0$ will play a crucial role. That is, in effect we will have to take into account the range restriction of l_z expressed by (5.1).

Both of these assumptions will only be used to ensure the validity of the Lipschitz change of variables in (5.31).

Let us begin with the proof. By interpolation with the case $p > 2$, it suffices to prove an estimate without decay; that is, it suffices to show that

$$\left\| \sum_{m > -l_z + \frac{v_z}{\alpha}} A_{u_z, l_z} P_m^{(1)} P_{-\alpha l_z + k}^{(2)} f(z) \right\|_{L_z^p} \lesssim \|f\|_p. \quad (5.11)$$

Again we get rid of the z -dependence of l_z by inserting a sum over all possible values of l_z , similar to (5.7). In particular, we bound the expression inside the L^p norm on the left hand side of (5.11) by⁷

$$\left(\sum_{l: v_z \in E_l} \left| \sum_{m > -l + \frac{v_z}{\alpha}} A_{u_z, l} P_m^{(1)} P_{-\alpha l + k}^{(2)} f(z) \right|^2 \right)^{1/2}. \quad (5.12)$$

We view the above expression as an l^2 -valued function. When estimating the L^p norm of it, we would like to appeal to certain interpolation for an l^q -valued function. However, for such a purpose, it is not convenient to carry the summation $\sum_{l: v_z \in E_l}$. It will be more convenient to see (5.12) as

$$\left(\sum_{l \in \mathbb{Z}} \left| \mathbb{1}_{E_l}(v_z) \sum_{m > -l + \frac{v_z}{\alpha}} A_{u_z, l} P_m^{(1)} P_{-\alpha l + k}^{(2)} f(z) \right|^2 \right)^{1/2}. \quad (5.13)$$

To estimate (5.13) we will run an interpolation argument which requires one to consider estimates of the form

$$\left\| \left(\sum_{l \in \mathbb{Z}} \left| \mathbb{1}_{E_l}(v_z) \sum_{m > -l + \frac{v_z}{\alpha}} A_{u_z, l} P_m^{(1)} P_{-\alpha l + k}^{(2)} f(z) \right|^q \right)^{1/q} \right\|_p \lesssim \left\| \left(\sum_{l \in \mathbb{Z}} \left| P_{-\alpha l + k}^{(2)} f \right|^q \right)^{1/q} \right\|_p. \quad (5.14)$$

When $q = \infty$, it is not difficult to see that the expression inside the L^p norm on the left hand side of the last display is bounded by

$$\mathcal{M}_{u, \epsilon_0}^{(\alpha)}(M_S f)(z). \quad (5.15)$$

Here M_S denotes the strong maximal operator. Recall that we already proved the L^p boundedness of $\mathcal{M}_{u, \epsilon_0}^{(\alpha)}$ for all $p > 2$ in Section 5.2. This implies

$$(5.14) \text{ holds for all } p > 2 \text{ and } q = \infty. \quad (5.16)$$

We will interpolate estimate (5.16) with

$$\left\| \mathbb{1}_{E_l}(v_z) \sum_{m > -l + \frac{v_z}{\alpha}} A_{u_z, l} P_m^{(1)} P_{-\alpha l + k}^{(2)} f(z) \right\|_p \lesssim \|f\|_p, \quad (5.17)$$

for all $p > 1$. Suppose for the moment that (5.17) holds. This gives

$$(5.14) \text{ holds for all } p = q > 1. \quad (5.18)$$

Then, interpolating (5.18) with (5.16) we obtain that

$$\|(5.13)\|_p \lesssim \|f\|_p, \quad (5.19)$$

for all $p > \frac{4}{3}$. Combining this with (5.15) implies

$$(5.14) \text{ holds for } p > \frac{4}{3} \text{ and } q = \infty. \quad (5.20)$$

We interpolate (5.20) and (5.17) to see that (5.19) holds for all $p > \frac{8}{7}$. Repeating this interpolation procedure sufficiently many times, we obtain that (5.13) is bounded on L^p for all $p > 1$. We learnt this interpolation trick from Nagel, Stein and Wainger [NSW78].

Before we turn to the proof of (5.17) we need to introduce some new notation. For $n \in \mathbb{Z}$ and $z \in \mathbb{R}^2$ we define

$$u_n(z) := \begin{cases} u_z & \text{if } v_z = n, \\ 2^n & \text{if } v_z \neq n. \end{cases}$$

⁷At this point we fix ϵ_0 possibly depending on α so that this estimate becomes valid.

Note carefully that $u_n(z)$ is different from $u_z^{(n)}$. Both take values in the interval $[2^n, 2^{n+1})$. However, the function u_n retains more of the regularity of u than $z \mapsto u_z^{(n)}$. This will be important during the proof of (5.17) (see (5.29), (5.38)).

Eliminating the z -dependence of v_z by introducing another sum we estimate the left hand side of (5.17) by

$$\left\| \left(\sum_{v \in E_l} \left| \sum_{m > -l + \frac{v}{\alpha}} A_{u_v(z), l} P_m^{(1)} P_{-\alpha l + k}^{(2)} f(z) \right|^2 \right)^{1/2} \right\|_p. \quad (5.21)$$

By the triangle inequality this is no greater than

$$\sum_{m > 0} \left\| \left(\sum_{v \in E_l} \left| A_{u_v(z), l} P_{m-l + \frac{v}{\alpha}}^{(1)} P_{-\alpha l + k}^{(2)} f(z) \right|^2 \right)^{1/2} \right\|_p. \quad (5.22)$$

Thus (5.17) will follow if we can show that there exist constants $\gamma_p > 0$ such that for every $k, m \in \mathbb{N}, l \in \mathbb{Z}$ we have

$$\left\| \left(\sum_{v \in E_l} \left| A_{u_v(z), l} P_{m-l + \frac{v}{\alpha}}^{(1)} P_{-\alpha l + k}^{(2)} f(z) \right|^2 \right)^{1/2} \right\|_p \lesssim 2^{-\gamma_p \max\{k, m\}} \|f\|_p. \quad (5.23)$$

We first prove (5.23) for $p = 2$. By Lemma 4.1 we obtain that for all $p > 2$ there exists $\gamma_p > 0$ such that

$$\left\| A_{u_v(z), l} P_{m-l + \frac{v}{\alpha}}^{(1)} P_{-\alpha l + k}^{(2)} f(z) \right\|_p \lesssim 2^{-\gamma_p \max\{k, m\}} \|f\|_p. \quad (5.24)$$

Thus, in order to show (5.23) for $p = 2$ it will be enough to show the estimate

$$\left\| A_{u_v(z), l} P_{m-l + \frac{v}{\alpha}}^{(1)} P_{-\alpha l + k}^{(2)} f(z) \right\|_p \lesssim \|f\|_p. \quad (5.25)$$

for an arbitrary $p < 2$. In the following we will prove something stronger, namely that

$$\left\| A_{u_v(z), l} f(z) \right\|_p \lesssim \|f\|_p \quad (5.26)$$

holds for all $p > 1$ provided that $v \in E_l$. Before commencing the proof of (5.26), we need to introduce another auxiliary function. If at a point $z = (x, y)$ we have $v_z = n$, then let $\tilde{u}_n(z) := u_z$. We now extend \tilde{u}_n to the whole space, requiring only that

$$\tilde{u}_n(z) \in [2^n, 2^{n+1}) \text{ and } \|\tilde{u}_n\|_{\text{Lip}} \leq 2\|u\|_{\text{Lip}}. \quad (5.27)$$

The advantage is that, unlike u_n , the function \tilde{u}_n is Lipschitz continuous with controlled Lipschitz norm. This will later be of key importance in ensuring the validity of the Lipschitz change of variables (5.31). Most importantly, the way that \tilde{u}_n and u_n are designed is such that they stay very “close” to each other, in the sense that if we define the following maximal operator:

$$\mathcal{M}_{\text{dyad}} f(z) := \sup_{j \in \mathbb{Z}} \sup_{\epsilon > 0} \frac{1}{2\epsilon} \int_{-\epsilon}^{\epsilon} |f(x - t, y - 2^j[t]^\alpha)| dt. \quad (5.28)$$

then we have the pointwise estimate

$$A_{u_v(z), l} f(z) \lesssim \mathcal{M}_{\text{dyad}} f(z) + A_{\tilde{u}_v(z), l} f(z) \quad (5.29)$$

for all $v, l \in \mathbb{Z}, z \in \mathbb{R}^2$. The maximal operator $\mathcal{M}_{\text{dyad}}$ is bounded in L^p for all $p > 1$.

Lemma 5.1. *For each $p > 1$, we have*

$$\|\mathcal{M}_{\text{dyad}} f\|_p \lesssim \|f\|_p. \quad (5.30)$$

The implicit constant depends only on p .

We postpone the proof of this fact until the end of this subsection and turn our attention toward the second term in (5.29). By Minkowski’s integral inequality, the L^p norm of $A_{\tilde{u}_v(z), l} f(z)$ is at most

$$\int_{\mathbb{R}} \left(\int_{\mathbb{R}^2} (f(x - t, y - \tilde{u}_v(z)[t]^\alpha) \psi_l(2^{\frac{v}{\alpha}} r_z t))^p dz \right)^{1/p} \frac{dt}{|t|}.$$

where $r_z := (2^{-v} \tilde{u}_n(z))^\beta \approx 1$. To bound this term, we apply the change of variables

$$\begin{cases} x_1 &:= x - t \\ y_1 &:= y - \tilde{u}_v(x, y) [t]^\alpha \end{cases} \quad (5.31)$$

This is the only point at which we use the Lipschitz regularity of u and it is also where we need to exploit the range restriction on l , i.e. that $2^l \leq c_\alpha \epsilon_0 2^{\frac{v}{\alpha}}$. Indeed, computing the determinant of the corresponding Jacobian J , we have

$$\det J = \left| \frac{\partial(x_1, y_1)}{\partial(x, y)} \right| = \left| \begin{array}{cc} 1 & 0 \\ -\frac{\partial \tilde{u}_v}{\partial x} [t]^\alpha & 1 - \frac{\partial \tilde{u}_v}{\partial y} [t]^\alpha \end{array} \right| = 1 - \frac{\partial \tilde{u}_v}{\partial y} [t]^\alpha. \quad (5.32)$$

Observe that t in (5.31) obeys

$$|[t]^\alpha| \approx 2^{l - \frac{v}{\alpha}} \leq c_\alpha \epsilon_0$$

and hence

$$|\det J| = \left| 1 - \frac{\partial \tilde{u}_v}{\partial y} [t]^\alpha \right| \geq 1 - \tilde{c}_\alpha \|\tilde{u}_v\|_{\text{Lip}} \epsilon_0 \geq \frac{1}{2}, \quad (5.33)$$

where we have chosen ϵ_0 to be smaller than $(2\tilde{c}_\alpha \|u\|_{\text{Lip}})^{-1}$. Here \tilde{c}_α is a constant depending only on α . This shows that the change of variables (5.31) is valid and therefore

$$\|A_{\tilde{u}_v(z), l} f\|_p \lesssim \|f\|_p \quad (5.34)$$

holds for all $p > 1$. Hence we can infer that also (5.26) holds. This concludes the proof of (5.23) in the case $p = 2$.

It remains to prove (5.23) for the remaining values of p . By interpolation with (5.23) at $p = 2$, it suffices to prove that we have the estimate without decay,

$$\left\| \left(\sum_{v \in E_l} \left| A_{u_v(z), l} P_{m-l+\frac{v}{\alpha}}^{(1)} P_{-\alpha l+k}^{(2)} f(z) \right|^2 \right)^{1/2} \right\|_p \lesssim \|P_{-\alpha l+k}^{(2)} f\|_p \sim \left\| \left(\sum_{v \in \mathbb{Z}} \left| P_{-l+m+\frac{v}{\alpha}}^{(1)} P_{-\alpha l+k}^{(2)} f \right|^2 \right)^{1/2} \right\|_p \quad (5.35)$$

for all $p > 1$. We will again use the interpolation trick from Nagel, Stein and Waigner [NSW78]. Thus we consider the more general estimate

$$\left\| \left(\sum_{v \in E_l} \left| A_{u_v(z), l} P_{m-l+\frac{v}{\alpha}}^{(1)} P_{-\alpha l+k}^{(2)} f(z) \right|^q \right)^{1/q} \right\|_p \lesssim \left\| \left(\sum_{v \in \mathbb{Z}} \left| P_{-l+m+\frac{v}{\alpha}}^{(1)} P_{-\alpha l+k}^{(2)} f \right|^q \right)^{1/q} \right\|_p \quad (5.36)$$

for $1 < p < \infty$ and $1 < q \leq \infty$.

Recall that in Section 5.2 we proved that

$$\|\mathcal{M}_{u, \epsilon_0}^{(\alpha)} f\|_p \lesssim \|f\|_p \text{ for all } p > 2. \quad (5.37)$$

Moreover, we have the pointwise bound

$$|A_{u_v(z), l} f(z)| \lesssim \mathcal{M}_{\text{dyad}} f(z) + \mathcal{M}_{u, \epsilon_0}^{(\alpha)} f(z) \text{ for all } v, l \in \mathbb{Z}. \quad (5.38)$$

Note that this pointwise estimate does not hold if u_v is replaced by \tilde{u}_v , because we do not know how \tilde{u}_v behaves outside of the region where it coincides with the original Lipschitz function u .

From (5.37), (5.38) and Lemma 5.1 we see that (5.36) holds for $q = \infty$ and $p > 2$. Moreover, by (5.26) we know that (5.36) holds for all $q = p > 1$. By interpolation we obtain (5.36) for $q = 2$ and $p > 4/3$. This in turn implies that

$$\|\mathcal{M}_{u, \epsilon_0}^{(\alpha)} f\|_p \lesssim \|f\|_p, \quad (5.39)$$

for all $p > 4/3$. Iterating this process sufficiently many times we prove (5.35) for every $p > 1$. This finishes the proof of (5.23) and thereby also concludes the proof of Theorem 1.1.

We close this subsection with several remarks regarding the last part of the proof:

- (1) In the proof of (5.26), we did not need the full strength of Lemma 5.1. For that purpose it would have been sufficient to a fixed j in the definition of $\mathcal{M}_{\text{dyad}}$ instead of taking the supremum over $j \in \mathbb{Z}$. The full strength of Lemma 5.1 is only needed during the interpolation procedure that is used for proving (5.23) for $p < 2$.
- (2) One might wonder why we did not use the auxiliary function \tilde{u}_v right away, rather than first introducing the auxiliary function u_v . Again, the point is that this would cause the interpolation argument for $p < 2$ to fail, because the pointwise estimate (5.38) would no longer be available.

5.4. Proof of Lemma 5.1. Before we start the proof of this lemma, we need to emphasize that this lemma was first established by Hong, Kim and Yang [HKY09]. Indeed, their results have a much wider scope, in the sense that they considered general polynomials in all dimensions. Due to such generality, their proof is significantly more intricate. For the sake of completeness, we provide here the proof of Lemma 5.1.

We will present the argument for $\alpha = 2$ and $[t]^2 = t^2$; the general case follows *mutatis mutandis*.

For $k, j \in \mathbb{Z}$ we set

$$M_k^j f(x, y) := \left| \frac{1}{2^k} \int_{2^k}^{2^{k+1}} f(x - t, y - 2^{-j} t^2) dt \right|. \quad (5.40)$$

Fix $j \in \mathbb{Z}$ consider the maximal operator along the parabola $(t, 2^{-j} t^2)$, which is given by

$$M^j f(x, y) := \sup_{k \in \mathbb{Z}} M_k^j f(x, y). \quad (5.41)$$

Hence

$$\mathcal{M}_{\text{dyad}} f(x, y) = \sup_j M^j f(x, y) = \sup_{k, j} M_k^j f(x, y). \quad (5.42)$$

A key idea to prove the L^2 bounds for $\mathcal{M}_{\text{dyad}}$ is to compare its components M_k^j with some suitable smoother versions \widetilde{M}_k^j and prove bounds for the corresponding square-function

$$\left(\sum_{j \in \mathbb{Z}} \sum_{k \in \mathbb{Z}} |M_k^j - \widetilde{M}_k^j|^2 \right)^{1/2}. \quad (5.43)$$

As a first attempt for finding a good candidate for \widetilde{M}_k^j one may consider the *linearized* model

$$\bar{M}_k^j f(x, y) := \frac{1}{2^k \cdot 3 \cdot 2^{2k}} \int_{2^k}^{2^{k+1}} \int_{2^{2k}}^{2^{2k+2}} f(x - t, y - 2^{-j} \tau) dt d\tau. \quad (5.44)$$

It turns out that while the operator $\left(\sum_{k \in \mathbb{Z}} |M_k^j - \bar{M}_k^j|^2 \right)^{1/2}$ is indeed bounded on L^2 , one does not have a good control over the double indexed sum in (5.43). For this reason, one needs to choose a variant that is closer in spirit to (5.40), which is simultaneously smoother and preserves the *quadratic* nature of the initial operator:

$$\widetilde{M}_k^j f(x, y) := \frac{1}{2^k \cdot 2^k} \int_{2^k}^{2^{k+1}} \int_{2^k}^{2^{k+1}} f(x - t, y - 2^{-j} \tau^2) dt d\tau. \quad (5.45)$$

It is not difficult to see that we have the pointwise bound

$$\widetilde{M}_k^j f(x, y) \lesssim M_S f(x, y), \quad (5.46)$$

and thus, in order to prove the L^p bound of $\mathcal{M}_{\text{dyad}}$, it suffices to prove the L^p bound for

$$\left(\sum_{j \in \mathbb{Z}} \sum_{k \in \mathbb{Z}} |M_k^j - \widetilde{M}_k^j|^2 \right)^{1/2}. \quad (5.47)$$

As opposed to the general L^p case, the proof for the L^2 boundedness of (5.47) is significantly simpler and thus we choose to present it first.

5.4.1. **L^2 boundedness.** In order to prove the L^2 bounds we rely on Plancherel's theorem. Naturally, we start by analyzing the multipliers for the corresponding operators M_k^j and \widetilde{M}_k^j .

The multiplier of M_k^j is given by

$$m_k^j(\xi, \eta) := \int_1^2 e^{i2^k t \xi + i2^{2k-j} t^2 \eta} dt, \quad (5.48)$$

while the multiplier of the operator (5.45) is given by

$$\widetilde{m}_k^j(\xi, \eta) := \int_1^2 \int_1^2 e^{i2^k t \xi + i2^{2k-j} \tau^2 \eta} dt d\tau. \quad (5.49)$$

By Plancherel's theorem, it suffices to prove that

$$\sum_{j \in \mathbb{Z}, k \in \mathbb{Z}} \left| m_k^j(\xi, \eta) - \widetilde{m}_k^j(\xi, \eta) \right|^2 \leq C, \quad (5.50)$$

for all $(\xi, \eta) \in \mathbb{R}^2$ and some universal constant $C > 0$. This in turn follows from

$$\sum_{j \in \mathbb{Z}, k \in \mathbb{Z}} \left| m_k^j(\xi, \eta) - \widetilde{m}_k^j(\xi, \eta) \right| \lesssim 1, \text{ for all } \xi \sim 1 \text{ and } \eta \sim 1. \quad (5.51)$$

- **Case I:** $k + 10 \geq j \geq 0$. Applying van der Corput's lemma, we have

$$\left| m_k^j(\xi, \eta) - \widetilde{m}_k^j(\xi, \eta) \right| \lesssim 2^{-(k-\frac{j}{2})}. \quad (5.52)$$

The last display is easily seen to be summable within the range $k \geq j \geq 0$.

- **Case II:** $4k \geq j > k + 10 \geq 0$ or $k \geq 0, j \leq 0$ or $j \leq 4k \leq 0$. For simplicity, we only detail the case $4k \geq j > k \geq 0$. Under this assumption, the phase functions in (5.49) and (5.48) do not admit any critical point. Thus,

$$\left| m_k^j(\xi, \eta) - \widetilde{m}_k^j(\xi, \eta) \right| \lesssim 2^{-k}. \quad (5.53)$$

By summing over $10 + k < j \leq 4k$, we obtain the upper bound $3k \cdot 2^{-k}$, which is summable in $k \in \mathbb{N}$.

- **Case III:** $j \geq 4k \geq 0$ or $k \leq 0, j \geq 0$ or $k \leq j \leq 0$. Again we only detail one case, that of $j \geq 4k \geq 0$. By definition, we have

$$\begin{aligned} m_k^j(\xi, \eta) - \widetilde{m}_k^j(\xi, \eta) &= \int_1^2 e^{i2^k t \xi + i2^{2k-j} t^2 \eta} dt - \int_1^2 \int_1^2 e^{i2^k t \xi + i2^{2k-j} \tau^2 \eta} dt d\tau \\ &= \int_1^2 e^{i2^k t \xi} \left(e^{i2^{2k-j} t^2 \eta} - 1 \right) dt - \int_1^2 \int_1^2 e^{i2^k t \xi} \left(e^{i2^{2k-j} \tau^2 \eta} - 1 \right) dt d\tau. \end{aligned} \quad (5.54)$$

By the fundamental theorem of calculus, the last display can be bounded by 2^{2k-j} , which is summable in $j \geq 4k \geq 0$.

Here we mention that the term $\bar{M}_k^j f$ from (5.44) would also work in this case.

- **Case IV:** $4k \leq j \leq k \leq 0$. In this case we will see the main difference between \bar{M}_k^j and \widetilde{M}_k^j . By definition,

$$\begin{aligned} m_k^j(\xi, \eta) - \widetilde{m}_k^j(\xi, \eta) &= \int_1^2 e^{i2^k t \xi + i2^{2k-j} t^2 \eta} dt - \int_1^2 \int_1^2 e^{i2^k t \xi + i2^{2k-j} \tau^2 \eta} dt d\tau \\ &= \int_1^2 \left(e^{i2^k \tau} - 1 \right) e^{i2^{2k-j} \tau^2 \eta} d\tau - \int_1^2 \int_1^2 \left(e^{i2^k t} - 1 \right) e^{i2^{2k-j} \tau^2 \eta} dt d\tau. \end{aligned} \quad (5.55)$$

By the fundamental theorem in calculus, we bound the last display by 2^k . Summing over $4k \leq j \leq k$, we obtain $|k| \cdot 2^k$, which is summable for $k \leq 0$.

This finishes the proof of the L^2 boundedness of our operator defined in (5.42).

5.4.2. L^p boundedness. In what follows we will make use of some ideas from Nagel, Stein and Wainger [NSW78] and Carlsson, Christ et al. [CCCD99]. We denote

$$M_k^j f := \mu_k^j * f, \text{ with } \widehat{\mu_k^j}(\xi, \eta) := m_k^j(\xi, \eta). \quad (5.56)$$

A key insight in [CCCD99], is to compare μ_k^j with σ_k^j , where

$$\sigma_k^j := \mu_k^j * [(\phi_k^j - \delta) \otimes (\tilde{\phi}_k^j - \delta)]. \quad (5.57)$$

Here ϕ and $\tilde{\phi}$ are two non-negative smooth functions supported on $[-1, 1]$ having mean one, while

$$\phi_k^j(t) := 2^{-k} \phi(2^{-k}t), \quad \tilde{\phi}_k^j(t) := 2^{-2k-2+j} \tilde{\phi}(2^{-2k-2+j}t), \quad (5.58)$$

and δ is the Dirac point mass at the origin. The meaning of the tensor product in (5.57) is that its first component acts on the first variable while its second component acts on the second variable.

The difference $\mu_k^j - \sigma_k^j$ can be bounded by the strong maximal operator, by noticing that the corresponding multiplier has fast decay at infinity. In particular,

$$\sup_{k,j} |(\mu_k^j - \sigma_k^j) * f| \lesssim M_S f. \quad (5.59)$$

Thus, it only remains to bound $\sup_{k,j} |\sigma_k^j * f|$. For this, we will perform a conical Littlewood-Paley decomposition for the function f :

$$P_k^{\text{Cone}} f(x, y) := \int_{\mathbb{R}} \hat{f}(\xi, \eta) \psi_k\left(\frac{\xi}{\eta}\right) e^{ix\xi + iy\eta} d\xi d\eta, \quad (5.60)$$

and write

$$\sup_{k,j} |\sigma_k^j * f| = \sup_{k,j} \left| \sigma_k^j * \left(\sum_{l \in \mathbb{Z}} P_{k-j+l}^{\text{Cone}} f \right) \right| \quad (5.61)$$

Remark 5.1. The case $l = 0$ corresponds to the case where the phase function of $\widehat{\mu_k^j} = m_k^j$ has a critical point.

To bound the term (5.61) on L^p , by the triangle inequality, it suffices to prove that

$$\left\| \sup_{k,j} |\sigma_k^j * (P_{k-j+l}^{\text{Cone}} f)| \right\|_p \lesssim 2^{-\gamma_p |l|} \|f\|_p, \quad (5.62)$$

for some $\gamma_p > 0$. This decay in l comes from the fact that away from the case $l = 0$ one never sees the critical point of the phase function of $\widehat{\mu_k^j}$.

In the following, we only focus on the case $l = 0$. The case of general $l \in \mathbb{Z}$ follows a similar approach, with the extra-twist of involving the non-stationary phase method (integration by parts). We refer to Carlsson, Christ et al. [CCCD99] for details.

Roughly speaking, estimate (5.62) follows from interpolating the L^2 bound of $\mathcal{M}_{\text{dyad}}$ with some simple endpoint bound. This interpolation is again in the spirit of Nagel, Stein and Wainger [NSW78]. To enable this, we need to rewrite (5.62) in a slightly different way:

$$\sup_{j,k} |\sigma_k^j * P_{k-j}^{\text{Cone}} f| = \sup_{j,k} |\sigma_k^{k-j} * P_j^{\text{Cone}} f| =: \sup_{j,k} |\tilde{\sigma}_k^j * P_j^{\text{Cone}} f|, \quad (5.63)$$

where

$$\tilde{\sigma}_k^j := \sigma_k^{k-j}. \quad (5.64)$$

We bound (5.63) by the square function:

$$\left(\sum_j \sup_k |\tilde{\sigma}_k^j * (P_j^{\text{Cone}} f)|^2 \right)^{1/2}. \quad (5.65)$$

To prove the L^p boundedness of the last expression, we first place it in a more general framework by considering inequalities of the form

$$\left\| \left(\sum_j \sup_k |\tilde{\sigma}_k^j * (P_j^{\text{Cone}} f)|^{q_1} \right)^{1/q_1} \right\|_{q_2} \lesssim \left\| \left(\sum_j |P_j^{\text{Cone}} f|^{q_1} \right)^{1/q_1} \right\|_{q_2}. \quad (5.66)$$

For the case $q_1 = \infty$ and $q_2 = 2$, by going back to $\sigma = \mu - (\mu - \sigma)$, it is not difficult to see that the left hand side of the last expression can be bounded by $\|\mathcal{M}_{\text{dyad}} f\|_2 + \|M_S f\|_2$, which, by the L^2 bound in the previous part, can be further bounded by $\|f\|_2$.

Hence, it remains to prove that for each fixed j and each $q_1 > 1$, one has

$$\left\| \sup_k |\tilde{\sigma}_k^j * (P_j^{\text{Cone}} f)| \right\|_{q_1} \lesssim \|P_j^{\text{Cone}} f\|_{q_1}. \quad (5.67)$$

That (5.65) is bounded on L^p for all $p > 1$ follows from an iterative interpolation argument in the spirit of [NSW78].

Now we prove (5.67).

We bound its left hand side by a square function, and prove that

$$\left\| \left(\sum_k |\tilde{\sigma}_k^j * (P_j^{\text{Cone}} f)|^2 \right)^{1/2} \right\|_{q_1} \lesssim \|P_j^{\text{Cone}} f\|_{q_1}. \quad (5.68)$$

By a simple anisotropic scaling, it suffices to consider the case $j = 0$.

Notice that working with the projection operator $P_0^{\text{Cone}} f$ means we are within a frequency cone having $\{(\xi, \eta) : \xi \sim \eta\}$. We continue by performing a finer frequency decomposition, as follows: denote by $P_{0,k}^{\text{Cone}}$ the frequency projection into the region $\xi \sim 2^k, \eta \sim 2^k$; then (5.68) is equivalent to

$$\left\| \left(\sum_k |\tilde{\sigma}_k^0 * \left(\sum_{l \in \mathbb{Z}} P_{0,k+l}^{\text{Cone}} f \right)|^2 \right)^{1/2} \right\|_{q_1} \lesssim \|P_0^{\text{Cone}} f\|_{q_1}. \quad (5.69)$$

By the triangle inequality, it suffices to prove for some $\lambda_{q_1} > 0$ that

$$\left\| \left(\sum_k |\tilde{\sigma}_k^0 * (P_{0,k+l}^{\text{Cone}} f)|^2 \right)^{1/2} \right\|_{q_1} \lesssim 2^{-\lambda_{q_1}|l|} \|P_0^{\text{Cone}} f\|_{q_1}. \quad (5.70)$$

The above estimate for $q_1 = 2$ follows simply from Plancherel's theorem and the mean zero property of $\tilde{\sigma}_k^0$. Thus, by interpolation and standard Littlewood-Paley theory, it suffices to prove that

$$\left\| \left(\sum_k |\tilde{\sigma}_k^0 * (P_{0,k+l}^{\text{Cone}} f)|^2 \right)^{1/2} \right\|_{q_1} \lesssim \left\| \left(\sum_k |P_{0,k+l}^{\text{Cone}} f|^2 \right)^{1/2} \right\|_{q_1}. \quad (5.71)$$

As before, we first consider this estimate in a more general framework of inequalities of the form

$$\left\| \left(\sum_k |\tilde{\sigma}_k^0 * (P_{0,k+l}^{\text{Cone}} f)|^{p_1} \right)^{1/p_1} \right\|_{q_1} \lesssim \left\| \left(\sum_k |P_{0,k+l}^{\text{Cone}} f|^{p_1} \right)^{1/p_1} \right\|_{q_1}. \quad (5.72)$$

The case $p_1 = \infty$ and $q_1 = 2$ follows from the L^2 boundedness of $\mathcal{M}_{\text{dyad}}$ while the case $p_1 = q_1 > 1$ is trivial. Applying the usual interpolation trick we obtain that (5.71) holds for all $q_1 > 1$. \square

APPENDIX A. A LOCAL SMOOTHING ESTIMATE

Let $\psi_0 : \mathbb{R} \rightarrow \mathbb{R}$ be a non-negative smooth function supported on $[1/2, 3]$ with $\psi_0(t) = 1$ for each $t \in [1, 2]$. Moreover, let $\varphi_0 : \mathbb{R} \rightarrow \mathbb{R}$ be a non-negative smooth function supported on $[-3, 3]$ with $\varphi_0(t) = 1$ for each $t \in [-1, 1]$. For a given positive real number u , let A_u denote the averaging operator

$$A_u f(x, y) := \int_{\mathbb{R}} f(x - ut, y - ut^\alpha) \psi_0(t) dt. \quad (\text{A.1})$$

Here α is a positive real number with $\alpha \neq 1$. For $k \in \mathbb{Z}$, let P_k denote a Littlewood-Paley projection operator on the plane, say

$$P_k f(x, y) := \int_{\mathbb{R}^2} e^{ix\xi + iy\eta} \hat{f}(\xi, \eta) \psi_0\left(\frac{1}{2^k}(\xi^2 + \eta^2)^{\frac{1}{2}}\right) d\xi d\eta. \quad (\text{A.2})$$

Then we have

Theorem A.1. *Let $k \in \mathbb{N}$ be a positive integer. For each positive $\alpha \neq 1$, and each $p > 2$, there exists $\gamma_{p,\alpha} > 0$ such that*

$$\| \sup_{u \in [1,2]} |A_u P_k f| \|_p \lesssim 2^{-\gamma_{p,\alpha} \cdot k} \|f\|_p. \quad (\text{A.3})$$

Here the implicit constant depends only on p and α .

In this section, we will prove Theorem A.1 by reducing it to a decoupling inequality for cones (see Proposition A.3) due to Bourgain [Bou13] and Bourgain and Demeter [BD15]. This follows the approach of Wolff [Wol00]. We will then provide a proof of the relevant decoupling inequality in the next section.

A.1. Several reductions. First of all, by applying the change of variable $t^\alpha \rightarrow s$, we see that it suffices to consider the case $\alpha > 1$. Secondly, to simplify our presentation, we will only work on the case $\alpha = 2$. The other values of $\alpha > 1$ can be handled in a similar way.

We take the Fourier transform of $A_u f$:

$$\widehat{A_u f}(\xi, \eta) = \hat{f}(\xi, \eta) \int_{\mathbb{R}} e^{iut\xi + iut^\alpha \eta} \psi_0(t) dt. \quad (\text{A.4})$$

By the stationary phase computation (for instance see page 360 in Stein [Ste93] or Lemma 1.2 in Iosevich [Ios94]),

$$\int_{\mathbb{R}} e^{it\xi + it^2 \eta} \psi_0(t) dt = a(\xi, \eta) e^{i\frac{\xi^2}{\eta}} + a_\infty(\xi, \eta). \quad (\text{A.5})$$

Here $a_\infty(\xi, \eta)$ is a smooth symbol, and $a(\xi, \eta)$ is a symbol belonging to the class $S^{-\frac{1}{2}}$ and is supported on

$$\{(\xi, \eta) : -10 \leq \frac{\xi}{\eta} \leq -\frac{1}{10}\}. \quad (\text{A.6})$$

The contribution from the smooth symbol $a_\infty(\xi, \eta)$ can be handled via a standard argument, see for instance Stein [Ste76]. We omit the details.

Now we turn to the former term on the right hand side of (A.5). Define

$$Tf(u, x, y) := \int_{\mathbb{R}^2} \hat{f}(\xi, \eta) a(u\xi, u\eta) e^{ix\xi + iy\eta + iu\frac{\xi^2}{\eta}} d\xi d\eta, \quad (\text{A.7})$$

and

$$T_k f(u, x, y) := \psi_0(u) \cdot T \circ P_k f(u, x, y). \quad (\text{A.8})$$

Moreover, we define a variant of T_k given by

$$S_k f(u, x, y) := \psi_0(u) \cdot \int_{\mathbb{R}^2} \widehat{P_k f}(\xi, \eta) \varphi_0\left(-\frac{\xi}{100\eta}\right) (1 + \xi^2 + \eta^2)^{-\frac{1}{4}} e^{ix\xi + iy\eta + iu\frac{\xi^2}{\eta}} d\xi d\eta. \quad (\text{A.9})$$

We will prove

Theorem A.2. *Let $k \in \mathbb{N}$ be a positive integer. For each $p > 2$, there exists $\gamma_p > 0$ such that*

$$\|S_k f\|_{L^p(\mathbb{R}^3)} \lesssim 2^{-(\frac{1}{p} + \gamma_p)k} \|P_k f\|_p. \quad (\text{A.10})$$

We will apply Theorem A.2 to prove

$$\|\Delta_u^{\frac{s}{2}} T_k f\|_{L^p(\mathbb{R}^3)} \lesssim 2^{-(\frac{1}{p} + \gamma_p)k + sk} \|f\|_p. \quad (\text{A.11})$$

Here

$$\Delta_u^{\frac{s}{2}} F(u, x, y) := \int_{\mathbb{R}} e^{iu\tau} (1 + |\tau|^2)^{\frac{s}{2}} \tilde{F}(\tau, x, y) d\tau, \quad (\text{A.12})$$

and \tilde{F} denotes the partial Fourier transform taken in the u variable only. The desired estimate (A.3) follows from (A.11) via the fractional L^∞ Sobolev embedding inequality

$$\|h\|_{L_u^\infty(\mathbb{R})} \lesssim \|\Delta_u^{\frac{s}{2}} h\|_{L_u^p(\mathbb{R})},$$

whenever $2 \leq p < \infty$ and $s > \frac{1}{p}$.

Now we come to the proof of (A.11). By taking the partial Fourier transform of $S_k f$ in the u variable, we see that

$$(S_k f)^\sim(\tau, x, y) = \int_{\mathbb{R}^2} \widehat{P_k f}(\xi, \eta) \varphi_0\left(-\frac{\xi}{100\eta}\right) (1 + \xi^2 + \eta^2)^{-\frac{1}{4}} \check{\psi}_0\left(\tau - \frac{\xi^2}{\eta}\right) e^{ix\xi + iy\eta} d\xi d\eta$$

is supported in the region $\{\tau \in \mathbb{R} : \tau \sim 2^k\}$. Hence by Young's inequality and Theorem A.2, we obtain

$$\|\Delta_u^{\frac{s}{2}} S_k f\|_{L^p(\mathbb{R}^3)} \lesssim 2^{sk} \|S_k f\|_{L^p(\mathbb{R}^3)} \lesssim 2^{-(\frac{1}{p} + \gamma_p)k + sk} \|f\|_p. \quad (\text{A.13})$$

The estimate (A.11) follows from (A.13) via the standard Hörmander-Mikhlin multiplier theorem, by realising that

$$\left| \partial_{\xi, \eta}^\alpha \frac{a(u\xi, u\eta)}{(1 + \xi^2 + \eta^2)^{-\frac{1}{4}}} \right| \lesssim_\alpha (1 + |\xi| + |\eta|)^{-|\alpha|} \text{ for all multi-indices } \alpha \in \mathbb{N}_0^2,$$

uniformly in $u \in [1, 2]$.

A.2. Proof of Theorem A.2 via a decoupling inequality. Define a rescaled version of the operator S_k from (A.9) by

$$E_k f(u, x, y) := \psi_k(u) \cdot \int_{\mathbb{R}^2} \widehat{P_0 f}(\xi, \eta) \varphi_0\left(-\frac{\xi}{100\eta}\right) (1 + \xi^2 + \eta^2)^{-\frac{1}{4}} e^{ix\xi + iy\eta + iu\frac{\xi^2}{\eta}} d\xi d\eta, \quad (\text{A.14})$$

where $\psi_k(u) := \psi_0(\frac{u}{2^k})$. By applying a change of variables

$$\xi \rightarrow 2^k \xi, \eta \rightarrow 2^k \eta \quad (\text{A.15})$$

to $S_k f$ and the desired estimate (A.10), we see that (A.10) is equivalent to

$$\|E_k f\|_{L^p(\mathbb{R}^3)} \lesssim 2^{\frac{k}{2} - \gamma_p \cdot k} \|f\|_p, \text{ for some } \gamma_p > 0. \quad (\text{A.16})$$

We apply a conical frequency decomposition for $P_0 f$. Let

$$\Sigma_k := \left\{ \frac{j}{2^{\frac{k}{2}}} : j \in \mathbb{Z}, -100 \cdot 2^{\frac{k}{2}} \leq j \leq 100 \cdot 2^{\frac{k}{2}} \right\} \quad (\text{A.17})$$

and decompose $P_0 f$ by writing

$$\widehat{P_0 f} = \sum_{\theta \in \Sigma_k} \widehat{f}_\theta := \sum_{\theta \in \Sigma_k} \psi_0\left(2^{\frac{k}{2}}\left(\frac{\xi}{\eta} - \theta\right)\right) \cdot \widehat{P_0 f}. \quad (\text{A.18})$$

The estimate (A.10) in Theorem A.2 follows immediately from the following two results.

Proposition A.3 (Bourgain [Bou13], Bourgain and Demeter [BD15]). *For each $2 \leq p \leq 4$ and each $\epsilon > 0$, we have*

$$\|E_k f\|_{L^p(\mathbb{R}^3)} \lesssim 2^{\frac{k}{2}(\frac{1}{2} - \frac{1}{p}) + \epsilon} \left(\sum_{\theta \in \Sigma_k} \|E_k f_\theta\|_{L^p(\mathbb{R}^3)}^p \right)^{\frac{1}{p}}. \quad (\text{A.19})$$

Here the implicit constant depends only on p , α and ϵ .

Lemma A.4. *For each $p \geq 2$, we have*

$$\left(\sum_{\theta \in \Sigma_k} \|E_k f_\theta\|_{L^p(\mathbb{R}^3)}^p \right)^{\frac{1}{p}} \lesssim 2^{\frac{k}{p}} \|f\|_{L^2(\mathbb{R}^2)}. \quad (\text{A.20})$$

The proof of Proposition A.3 can be found in Bourgain [Bou13] and the last section of Bourgain and Demeter [BD15]. Note that here we only rely on a weak form of the decoupling inequality; that is, the exponent p is in the restricted range $2 \leq p \leq 4$, but not $4 < p \leq 6$. The latter is part of the main content in Bourgain and Demeter [BD15]. As the proof of Proposition A.3 is short and can be made (essentially) self-contained, we will provide it in the next section.

The proof of Lemma A.4 is standard. Here we sketch the proof. By interpolation, it suffices to prove (A.20) for $p = 2$ and $p = \infty$. At $p = 2$, the proof is via a simple almost orthogonality argument. For each fixed $u \in [0, 2^k]$, we have

$$\left(\sum_{\theta \in \Sigma_k} \|E_k f_\theta(u, \cdot)\|_{L^2_{x,y}(\mathbb{R}^2)}^2 \right)^{\frac{1}{2}} \lesssim \left(\sum_{\theta \in \Sigma_k} \|f_\theta\|_{L^2(\mathbb{R}^2)}^2 \right)^{\frac{1}{2}} \lesssim \|f\|_{L^2(\mathbb{R}^2)}. \quad (\text{A.21})$$

In the end, we integrate in u and collect the factor $2^{\frac{k}{2}}$.

At $p = \infty$, by Young's inequality, it suffices to show that

$$\left\| \int_{\mathbb{R}^2} \varphi_0\left(-\frac{\xi}{100\eta}\right) \psi_0(\eta) \psi_0\left(2^{\frac{k}{2}}\left(\frac{\xi}{\eta} - \theta\right)\right) e^{ix\xi + iy\eta + iu\frac{\xi^2}{\eta}} d\xi d\eta \right\|_{L^1_{x,y}(\mathbb{R}^2)} \lesssim 1, \quad (\text{A.22})$$

uniformly in θ and $u \in [0, 2^k]$. Without loss of generality, we take $\theta = 0$. By the change of variable $2^{\frac{k}{2}}\xi \rightarrow \xi$, the last estimate is equivalent to

$$\left\| \int_{\mathbb{R}^2} \psi_0(\eta) \psi_0\left(\frac{\xi}{\eta}\right) e^{ix\xi + iy\eta + iu\frac{\xi^2}{\eta}} d\xi d\eta \right\|_{L^1_{x,y}(\mathbb{R}^2)} \lesssim 1, \quad (\text{A.23})$$

uniformly in $u \in [0, 1]$. However, by the non-stationary phase method, when $|x| + |y| \gg 1$, we always have

$$\left| \int_{\mathbb{R}^2} \psi_0(\eta) \psi_0\left(\frac{\xi}{\eta}\right) e^{ix\xi + iy\eta + iu\frac{\xi^2}{\eta}} d\xi d\eta \right| \lesssim \frac{1}{|x|^{10} + |y|^{10}}.$$

This further implies the desired estimate (A.23).

APPENDIX B. THE PROOF OF A DECOUPLING INEQUALITY

For a dyadic interval $\Delta \subset [0, 1]$, define the extension operator associated with Δ and the parabola (ξ, ξ^2) by

$$E_\Delta g(x) := \int_\Delta g(\xi) e^{ix_1\xi + ix_2\xi^2} d\xi. \quad (\text{B.1})$$

In this section we will prove

Theorem B.1 (Bourgain [Bou13]). *For each $2 \leq p \leq 4$ and $\epsilon > 0$, we have*

$$\|E_{[0,1]}g\|_{L^p(\mathbb{R}^2)} \lesssim \delta^{-(\frac{1}{2} - \frac{1}{p} + \epsilon)} \left(\sum_{\Delta \subset [0,1]; l(\Delta)=\delta} \|E_\Delta g\|_{L^p(\mathbb{R}^2)}^p \right)^{\frac{1}{p}}. \quad (\text{B.2})$$

Proposition A.3 follows from Theorem B.1 via Fubini's theorem and an iteration argument. This iteration first appeared in the work of Pramanik and Seeger [PS07]. We refer to Proposition 8.1 in Bourgain and Demeter [BD15] for the details.

In the remaining part, we will provide a proof of Theorem B.1. First of all, by a simple localisation argument, and by Hölder's inequality, the estimate (B.2) follows from

$$\|E_{[0,1]}g\|_{L^p(w_B)} \lesssim \delta^{-\epsilon} \left(\sum_{\Delta \subset [0,1]; l(\Delta)=\delta} \|E_\Delta g\|_{L^p(w_B)}^2 \right)^{\frac{1}{2}}, \quad (\text{B.3})$$

for each ball B of radius δ^{-2} . Here w_B is a weight associated with B given by

$$w_B(x) := \left(1 + \frac{\|x - c_B\|}{\delta^{-2}}\right)^{-N},$$

for a large integer N which will not be specified. Secondly, to prove (B.3) for all $2 \leq p \leq 4$, by interpolation with the trivial bound at $p = 2$, it suffices to look at the case $p = 4$. We refer to Garrigós and Seeger [GS10] for the details of such an interpolation argument.

The proof of (B.3) for $p = 4$ will be accomplished in three steps, which correspond to the following three subsections.

B.1. A bilinear restriction estimate for parabola. The following proposition follows simply via Plancherel's theorem.

Proposition B.2. *Let $R_1, R_2 \subset [0, 1]$ be two dyadic intervals with $\text{dist}(R_1, R_2) \geq \nu$ for some $\nu > 0$. We have the bilinear restriction estimate*

$$\left\| |E_{R_1} g_1 E_{R_2} g_2|^{\frac{1}{2}} \right\|_{L^4(\mathbb{R}^2)} \lesssim_{\nu} \|g_1\|_2^{\frac{1}{2}} \|g_2\|_2^{\frac{1}{2}}. \quad (\text{B.4})$$

The details are left to the interested reader.

B.2. A bilinear decoupling inequality. Recall that R_1 and R_2 are two dyadic intervals whose distance is not smaller than ν .

Proposition B.3. *We have a bilinear version of the desired decoupling inequality (B.3):*

$$\left\| (E_{R_1} g_1 E_{R_2} g_2)^{\frac{1}{2}} \right\|_{L^4(w_B)} \lesssim_{\nu} \left(\prod_{j=1}^2 \sum_{\Delta \subset R_j: l(\Delta)=\delta} \|E_{\Delta} g_j\|_{L^4(w_B)}^2 \right)^{\frac{1}{4}}, \quad (\text{B.5})$$

for each ball B of radius δ^{-2} .

We start by introducing some notation. Let τ_j be the δ^2 -neighbourhood of the parabola that lies on top of R_j ; that is,

$$\tau_j := \{(\xi, \xi^2 + \eta) : \xi \in R_j, |\eta| \leq \delta^2\}.$$

We let \mathcal{P}_j be finitely overlapping cover of τ_j with curved regions θ of the form

$$\theta = \{(\xi, \xi^2 + \eta) : \xi \in [c, c + \delta], |\eta| \leq \delta^2\} \text{ for some constant } c.$$

For a function f supported on τ_j , we let f_{θ} to denote the restriction of f on θ .

Let f_1 and f_2 be two functions supported on τ_1 and τ_2 separately. The bilinear estimate (B.4) implies

$$\|(\hat{f}_1 \hat{f}_2)^{\frac{1}{2}}\|_{L^4(w_B)} \lesssim \delta^{-1} \|f_1\|_{L^2(\mathbb{R}^2)}^{\frac{1}{2}} \|f_2\|_{L^2(\mathbb{R}^2)}^{\frac{1}{2}}. \quad (\text{B.6})$$

This can be done by applying (B.4) to each fibre $\{(\xi, \xi^2 + \eta_j)\}$ for fixed η_j ($j = 1, 2$), and then applying Minkowski's inequality to the variables η_j . Now we apply the L^2 orthogonality and a simple localisation argument, to further obtain

$$\|(\hat{f}_1 \hat{f}_2)^{\frac{1}{2}}\|_{L^4(w_B)} \lesssim \delta^{-1} \left(\prod_{j=1}^2 \sum_{\theta \in \mathcal{P}_j} \|\widehat{f_{j,\theta}}\|_{L^2(w_B)}^2 \right)^{\frac{1}{4}}. \quad (\text{B.7})$$

In the end, we apply Hölder's inequality to the right hand side of the last expression

$$\|(\hat{f}_1 \hat{f}_2)^{\frac{1}{2}}\|_{L^4(w_B)} \lesssim \left(\prod_{j=1}^2 \sum_{\theta \in \mathcal{P}_j} \|\widehat{f_{j,\theta}}\|_{L^4(w_B)}^2 \right)^{\frac{1}{4}}. \quad (\text{B.8})$$

This implies the desired estimate in Proposition B.3 by taking $f_j = E_{R_j} g_j$.

B.3. Bilinear decoupling implies linear decoupling. We come to the final step of proving the desired decoupling inequality (B.3). The idea is that the bilinear decoupling inequality in Proposition B.3 will imply (B.3). This is done via a simple version of the Bourgain-Guth argument from [BG11].

We proceed with the details. Fix a large constant $K \ll \delta^{-1}$. We split the interval $[0, 1]$ into smaller intervals of length K^{-1} . We use α to denote such an interval. Then on each ball B_K of radius K , the function $|E_\alpha g|$ behaves like a constant. Hence for convenience, we will use $|E_\alpha g|(B_K)$ to denote such a value.

For each given B_K , we let α^* denote the interval that maximises $|E_{\alpha^*} g|(B_K)$. We look at the collection of α , with $\text{dist}(\alpha^*, \alpha) \geq \frac{1}{K}$, and

$$|E_\alpha g|(B_K) \geq \frac{1}{10K} |E_{\alpha^*} g|(B_K).$$

There are two cases. The first case is that this collection is empty. Then

$$|E_{[0,1]} g|(x) \lesssim |E_{\alpha^*} g|(B_K), \text{ for each } x \in B_K. \quad (\text{B.9})$$

The second case is that this collection contains at least one element. Call it α^{**} . Then

$$|E_{[0,1]} g|(B_K) \lesssim K^3 |E_{\alpha^*} g|^{\frac{1}{2}}(B_K) |E_{\alpha^{**}} g|^{\frac{1}{2}}(B_K). \quad (\text{B.10})$$

Putting these two estimates together, we obtain

$$\begin{aligned} \|E_{[0,1]} g\|_{L^4(w_{B_K})} &\leq C \left(\sum_{\alpha: l(\alpha)=\frac{1}{K}} \|E_\alpha g\|_{L^4(w_{B_K})}^2 \right)^{\frac{1}{2}} \\ &\quad + C \cdot K^3 \left(\sum_{\text{dist}(\alpha_1, \alpha_2) \geq \frac{1}{K}} \left\| |E_{\alpha_1} g|^{\frac{1}{2}} |E_{\alpha_2} g|^{\frac{1}{2}} \right\|_{L^4(w_{B_K})}^2 \right)^{\frac{1}{2}}, \end{aligned} \quad (\text{B.11})$$

for a universal constant C . We raise both sides of this estimate to the 4-th power, and sum over all balls B_K inside B , a ball of radius δ^{-2} , to obtain that (B.11) indeed holds true with B_K being replaced by B . Now we apply the bilinear decoupling inequality that has been proven in the previous subsection, to obtain

$$\|E_{[0,1]} g\|_{L^4(w_B)} \leq C \left(\sum_{\alpha: l(\alpha)=\frac{1}{K}} \|E_\alpha g\|_{L^4(w_B)}^2 \right)^{\frac{1}{2}} + C_K \cdot K^{10} \left(\sum_{\Delta: l(\Delta)=\delta} \|E_\Delta g\|_{L^4(w_B)}^2 \right)^{\frac{1}{2}}, \quad (\text{B.12})$$

for a possibly larger C and a constant C_K depending on K . In the end, by invoking a parabolic rescaling, we iterate the last estimate for $\log_K(\frac{1}{\delta})$ many times, and obtain

$$\|E_{[0,1]} g\|_{L^p(w_B)} \leq C^{\log_K(\frac{1}{\delta})} \cdot C_K \cdot K^{10} \left(\sum_{\Delta: l(\Delta)=\delta} \|E_\Delta g\|_{L^p(w_B)}^2 \right)^{\frac{1}{2}}. \quad (\text{B.13})$$

We just need to observe that by choosing K large enough, compared with C , we can always have the desired estimate (B.3).

REFERENCES

- [Bat13] Bateman, M. *Single annulus L^p estimates for Hilbert transforms along vector fields*. Rev. Mat. Iberoam. 29 (2013), no. 3, 1021-1069.
- [BT13] Bateman, M. and Thiele, C. *L^p estimates for the Hilbert transforms along a one-variable vector field*. Anal. PDE 6 (2013), no. 7, 1577-1600.
- [Bou86] Bourgain, J. *Averages in the plane over convex curves and maximal operators*. J. Analyse Math. 47 (1986), 69-85.
- [Bou89] Bourgain, J. *A remark on the maximal function associated to an analytic vector field*. Analysis at Urbana, Vol. I (Urbana, IL, 1986-1987), 111-132, London Math. Soc. Lecture Note Ser., 137, Cambridge Univ. Press, Cambridge, 1989.
- [Bou13] Bourgain, J. *Moment inequalities for trigonometric polynomials with spectrum in curved hypersurfaces*. Israel J. Math. 193 (2013), no. 1, 441-458.

- [BD15] Bourgain, J. and Demeter, C. *The proof of the l^2 decoupling conjecture*. Ann. of Math. (2) 182 (2015), no. 1, 351-389.
- [BG11] Bourgain, J. and Guth, L. *Bounds on oscillatory integral operators based on multilinear estimates*. Geom. Funct. Anal. 21 (2011), no. 6, 1239-1295.
- [CSWW99] Carbery, A., Seeger, A., Wainger, S. and Wright, J. *Classes of singular integral operators along variable lines*. J. Geom. Anal. 9 (1999), no. 4, 583-605.
- [CCCD99] Carlsson, H., Christ, M., Cordoba, A., Duoandikoetxea, J., Rubio de Francia, J. L., Vance, J., Wainger, S. and Weinberg, D. *L^p estimates for maximal functions and Hilbert transforms along flat convex curves in R^2* . Bull. Amer. Math. Soc. (N.S.) 14 (1986), no. 2, 263-267.
- [CNSW99] Christ, M., Nagel, A., Stein, E. and Wainger, S. *Singular and maximal Radon transforms: analysis and geometry*. Ann. of Math. (2) 150 (1999), no. 2, 489-577.
- [FS71] Fefferman, C. and Stein, E. *Some Maximal Inequalities*. American Journal of Mathematics. Vol. 93, No. 1 (Jan., 1971), pp. 107-115.
- [GS10] Garrigós, G. and Seeger, A. *A mixed norm variant of Wolff's inequality for paraboloids*. Harmonic analysis and partial differential equations, 179-197, Contemp. Math., 505, Amer. Math. Soc., Providence, RI, 2010.
- [Guo16] Guo, S. *Single annulus estimates for the variation-norm Hilbert transforms along Lipschitz vector fields*. To appear in the Proc. Amer. Math. Soc.
- [GPRY16] Guo, S., Pierce, L., Roos, J. and Yung, P.-L. *Polynomial Carleson operators along monomial curves in the plane*. arXiv:1605.05812
- [HKY09] Hong, S., Kim, J. and Yang, C. *Maximal operators associated with vector polynomials of lacunary coefficients*. Canad. J. Math. 61 (2009), no. 4, 807-827.
- [Ios94] Iosevich, A. *Maximal operators associated to families of flat curves in the plane*. Duke Math. J. 76 (1994), no. 2, 633-644.
- [Kar07] Karagulyan, G. *On unboundedness of maximal operators for directional Hilbert transforms*. Proc. Amer. Math. Soc. 135 (2007), no. 10, 3133-3141 (electronic).
- [LL06] Lacey, M. and Li, X. *Maximal theorems for the directional Hilbert transform on the plane*. Trans. Amer. Math. Soc. 358 (2006), no. 9, 4099-4117.
- [LL10] Lacey, M. and Li, X. *On a conjecture of E. M. Stein on the Hilbert transform on vector fields*. Mem. Amer. Math. Soc. 205 (2010), no. 965, viii+72 pp.
- [Lie09] Lie, V. *The (weak- L^2) boundedness of the quadratic Carleson operator*. Geom. Funct. Anal. 19 (2009), no. 2, 457-497.
- [Lie11] Lie, V. *The Polynomial Carleson Operator*. arXiv:1105.4504.
- [Lie15] Lie, V. *The boundedness of the Bilinear Hilbert Transform along non-flat smooth curves*. Amer. J. Math., 137(2), pp. 313-364, 2015.
- [Lie16] Lie, V. *On the Boundedness of the Bilinear Hilbert Transform along non-flat smooth curves. The Banach Triangle case (L^r , $1 \leq r < \infty$)*. Rev. Mat. Iberoam, to appear 2016.
- [MR98] Marletta, G. and Ricci, F. *Two-parameter maximal functions associated with homogeneous surfaces in R^n* . Studia Math. 130 (1998), no. 1, 53-65.
- [MSS92] Mockenhaupt, G., Seeger, A. and Sogge, C. *Wave front sets, local smoothing and Bourgain's circular maximal theorem*. Ann. of Math. (2) 136 (1992), no. 1, 207-218.
- [Mus14] Muscalu, C. *Calderon commutators and the Cauchy integral on Lipschitz curves revisited I. First commutator and generalizations*. Rev. Mat. Iberoam. 30 (2014), no. 2, 727-750.
- [NSW78] Nagel, A., Stein, E. and Wainger, S. *Differentiation in lacunary directions*. Proc. Nat. Acad. Sci. U.S.A. 75 (1978), no. 3, 1060-1062.
- [NSW79] Nagel, A., Stein, E. and Wainger, S. *Hilbert transforms and maximal functions related to variable curves*. Harmonic analysis in Euclidean spaces (Proc. Sympos. Pure Math., Williams Coll., Williamstown, Mass., 1978), Part 1, pp. 95-98, Proc. Sympos. Pure Math., XXXV, Part, Amer. Math. Soc., Providence, R.I., 1979.
- [OSTTW12] Oberlin, R., Seeger, A., Tao, T., Thiele, C. and Wright, J. *A variation norm Carleson theorem*. J. Eur. Math. Soc. (JEMS) 14 (2012), no. 2, 421-464.
- [PS07] Pramanik, M. and Seeger, A. *L^p regularity of averages over curves and bounds for associated maximal operators*. Amer. J. Math. 129 (2007), no. 1, 61-103.
- [See94] Seeger, A. *L^2 -estimates for a class of singular oscillatory integrals*. Math. Res. Lett. 1 (1994), no. 1, 65-73.
- [Ste76] Stein, E. *Maximal functions. I. Spherical means*. Proc. Nat. Acad. Sci. U.S.A. 73 (1976), no. 7, 2174-2175.
- [Ste93] Stein, E. *Harmonic analysis: real-variable methods, orthogonality, and oscillatory integrals. With the assistance of Timothy S. Murphy*. Princeton Mathematical Series, 43. Monographs in Harmonic Analysis, III. Princeton University Press, Princeton, NJ, 1993. xiv+695 pp.
- [SS12] Stein, E. and Street, B. *Multi-parameter singular Radon transforms III: Real analytic surfaces*. Adv. Math. 229 (2012), no. 4, 2210-2238.

- [SW01] Stein, E. and Wainger, S. *Oscillatory integrals related to Carleson's theorem*. Math. Res. Lett. 8 (2001), no. 5-6, 789-800.
- [SW03] Seeger, A. and Wainger, S. *Singular Radon transforms and maximal functions under convexity assumptions*. Rev. Mat. Iberoamericana 19 (2003), no. 3, 1019-1044.
- [Wol00] Wolff, T. *Local smoothing type estimates on L^p for large p* . Geom. Funct. Anal. 10 (2000), no. 5, 1237-1288.

SHAOMING GUO: UNIVERSITY OF BONN, ENDENICHER ALLEE 60, 53115 BONN, GERMANY. CURRENT ADDRESS: INDIANA UNIVERSITY BLOOMINGTON, 831 E THIRD ST, BLOOMINGTON, IN 47405, USA
E-mail address: `shaoguo@iu.edu`

JONATHAN HICKMAN: UNIVERSITY OF CHICAGO, 5734 S. UNIVERSITY AVENUE, CHICAGO, ILLINOIS, 60637.
E-mail address: `jehickman@uchicago.edu`

VICTOR LIE: DEPARTMENT OF MATHEMATICS, PURDUE, IN 46907 USA AND INSTITUTE OF MATHEMATICS OF THE ROMANIAN ACADEMY, BUCHAREST, RO 70700, P.O. Box 1-764, ROMANIA.
E-mail address: `vlie@purdue.edu`

JORIS ROOS: UNIVERSITY OF BONN, MATHEMATICAL INSTITUTE, ENDENICHER ALLEE 60, 53115 BONN, GERMANY
E-mail address: `jroos@math.uni-bonn.de`